BULLETIN

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

EARLE R. HEDRICK
DAVID R. CURTISS WILLIAM R. LONGLEY

WITH THE ASSISTANCE OF

HAROLD T. DAVIS HARRY W. KUHN
THEOPHIL H. HILDEBRANDT DERRICK N. LEHMER
JOHN R. KLINE HENRY L. RIETZ
CAROLINE E. SEELY

VOLUME XL, NUMBER 8

AUGUST, 1934

PUBLISHED BY THE SOCIETY
MENASHA, WIS., AND NEW YORK
1934

Whole No. 419 $9.00 a Year

PUBLISHED MONTHLY

Entered as second class matter April 24, 1926 at the post office at Menasha, Wis., under the act of August 24, 1912.

PERIODICAL PUBLICATIONS OF THE AMERICAN MATHEMATICAL SOCIETY

BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY

This journal is published monthly. It is the official organ of the Society. One of its functions is to report official acts of the Society and the details of its meetings, with abstracts of all research papers presented. In addition to these official reports, it publishes, also, short original papers on mathematical topics, and some of the longer officially invited addresses presented before the Society. These papers form the body of the Bulletin. It contains also announcements and notes on matters of current interest, and reviews of advanced mathematical books.

Price per annual volume, $9.00. Sets of early volumes are available as follows: vols. 12-36 (1905-1930), $7.00 each; vols. 37 to date, $9.00 each. Vols. 1-11 and the three volumes of the Bulletin of the New York Mathematical Society (the precursor of the present Bulletin) are out of print and scarce; sets can sometimes be obtained through the Society's New York office.

TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY

This journal is devoted to research in pure and applied mathematics.

Beginning with 1935, two volumes of three numbers each of the Transactions will be published per year. The price per volume will be $6.00, so that the annual subscription price (six issues) will be $12.00. The Transactions has appeared as a quarterly since 1900, with the exception of the year 1922, when two volumes, of four numbers each, were published (vols. 23-24).

Sets of early volumes are available as follows: vols. 1-21 (1900-1920), $7.00 each; vols. 22-32 (1921-1930), $8.00 each; vols. 33-36 (1931-1934), $9.00 each.

Subscriptions and orders for the Bulletin or Transactions should be addressed to the American Mathematical Society, Menasha, Wisconsin, or 501 West 116th St., New York City.

THE ASYMPTOTIC SOLUTIONS OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS OF THE SECOND ORDER, WITH SPECIAL REFERENCE TO THE STOKES PHENOMENON†

BY R. E. LANGER

1. *Introduction.* From the title of my address you will already have observed that it is my intention to discuss with you the asymptotic representation of the solutions of certain differential equations, and that my point in doing this is largely to dwell upon the Stokes phenomenon which arises in that theory. There are at least two good reasons why I have chosen to do this. Firstly, the subject seems to me to have mathematical features of distinct interest and perhaps of importance, and, secondly, it has risen during the last dozen years to a position of importance in theoretical physics, both of the more classical type and of the modern quantum mechanics. The instance is indeed one in which the development of mathematics has not merely been hastened by the spur of the needs of growing physical theory, but in which investigators primarily or at least prominently identified as physicists have turned aside to mathematical deductions to fill their own needs.

The Stokes phenomenon is named after its discoverer, the British mathematician and physicist, Sir G. G. Stokes. He was confronted by its manifestation in his study of the Bessel functions, and that it held puzzlement and difficulty for him he was not loath to acknowledge. With your permission I should, indeed, like to read you an abstract from a letter which he wrote at the time, as you will conclude, to a certain young lady:‡ London, March 19/57. "When the cat's away the mice may play. You are the cat and I am the poor little mouse. I have been doing what I guess you won't let me do when we are married, sitting up till 3 o'clock in the morning fighting hard against a mathematical difficulty. Some years ago I attacked an integral of Airy's, and after a severe trial reduced it to a readily calculable form. But there was one difficulty about it which,

† A symposium lecture delivered by invitation of the program committee at the meeting of the Society at Chicago, April 6, 1934.

‡ Sir George Gabriel Stokes, Memoirs and Scientific Correspondence, Cambridge, 1907, vol. 1, p. 62.

though I tried till I almost made myself ill, I could not get over, and at last I had to give it up and profess myself unable to master it. I took it up again a few days ago, and after two or three days' fight, the last of which I sat up till 3, I at last mastered it. · · · " A decade later the phenomenon was observed, again in connection with the Bessel functions, by Hankel, who speaks of his investigations as having led to "peculiar results which might perhaps serve as a basis for a general theory of semi-convergent series," those remarkable developments whose true nature was almost entirely unknown.†

2. *The Sturm-Liouville Equation.* Let me, to begin with, write down the partial differential equation

$$\frac{\partial}{\partial x}\left\{k(x)\frac{\partial T}{\partial x}\right\} - q(x)T = p(x)\frac{\partial T}{\partial t}$$

which controls the temperature T in the flow of heat along a slender rod in the direction x with time t. The coefficients k, p, and q are all essentially positive functions, determined by the physical characteristics of the rod and the surrounding medium. A familiar mode of procedure toward solving the equation is to substitute

$$T = f(t)u(x),$$

whereupon the equation decomposes into the form

$$\frac{f'(t)}{f(t)} = \frac{\dfrac{d}{dx}\left(k\dfrac{du}{dx}\right) - qu}{pu}.$$

In this the left member is a function of t alone and the right member is a function of x alone, and the equality therefore implies that each is independent of both x and t, say is equal to a quantity $-\lambda^2$. I am not interested in the equation in t. The equation for the component $u(x)$ is, however, of the form

$$\frac{d}{dx}\left\{k(x)\frac{du}{dx}\right\} + \{\lambda^2 p(x) - q(x)\}u = 0, \tag{1}$$

† H. Hankel, *Die Cylinderfunktionen erster und zweiter Art*, Mathematische Annalen, vol. 1 (1869), p. 467.

in which λ^2 is independent of x, but may be regarded as a parameter variable in its own right. The equation (1) is known as the differential equation of the Sturm-Liouville type, in recognition of the fundamental work done in connection with it by Sturm and Liouville about a century ago. It arises in the mathematical formulation of classical physical problems of such great diversity as the flow of heat, the vibrations of strings and membranes, the oscillations of rods and beams, the diffusion of substances through each other, the distribution of electrical potentials and current densities, the turbulent motions of fluids, the tides, etc. Mathematically it is the type equation which includes as special cases the differential equations of such standard functions as the trigonometric functions, the Bessel functions, the Legendre functions, the polynomials of Hermite, Tchebycheff, and Laguerre, certain confluent hypergeometric functions, the functions of Mathieu, and many others. Much more recently it has come into renewed prominence among the physicists as a type of the Schrödinger wave equation in the theory of quantum mechanics.

It is natural under these circumstances that much study has been given to the equation in the course of a century, with the result that in its present form its theory includes among others such beautiful chapters as those on the oscillation and comparison of solutions, the boundary problems, the expansion of arbitrary functions in series of solutions, the asymptotic forms of the solutions for large complex values of the variable, and the asymptotic dependence of the solutions upon the parameter λ. It is to the last two of these, and more particularly to the last one, to which I shall confine my considerations.

If in the form (1) the coefficient $k(x)$ is not zero on a given interval, and admits a second derivative, a familiar change of dependent variable reduces the equation to the form

$$u'' + \{\lambda^2\phi^2(x) - \chi(x)\}u = 0. \tag{2}$$

In this ϕ^2 and λ^2 are written in place of ϕ and λ only as a matter of convenience. There is no presumption that these quantities are positive; indeed, for many considerations they need not even be real. The form (2) has the advantage of great simplicity and will, therefore, be made the immediate basis of our discussion.

To avoid a dissipation of time upon matters which are after all of rather a secondary interest, I shall assume henceforth that where the equation (2) is to be considered its coefficients are continuous and, in fact, differentiable to whatever order may be required. To keep the discussion within proper bounds I shall have to confine myself very largely to real values of x and λ, and when doing so I shall suppose λ positive and shall refer to the range of values which x may take on as the interval $a \leqq x \leqq b$. It is to be understood, however, that unless the contrary is stated the possibilities $a = -\infty$, or $b = +\infty$, or both, are not excluded.

3. *The Asymptotic Solutions when $\phi^2(x)$ is Bounded From Zero.* If $\phi^2(x)$ is real and is bounded from zero on (a, b), we may define the root $\phi(x)$ so that $\phi/|\phi|$ is either 1 or i according as ϕ^2 is positive or negative. The change of variables

$$t = \frac{1}{\kappa}\int_*^x |\phi|\,dx, \qquad v = |\phi(x)|^{1/2}u, \qquad i\rho = \kappa\lambda\left\{\frac{\phi}{|\phi|}\right\},$$

in which κ and $*$ may be any constants with respect to x, reduces the differential equation to the simple form

$$\frac{d^2v}{dt^2} - \{\rho^2 + \omega_1(t)\}v = 0, \tag{3}$$

with the variable t ranging over an interval (t_a, t_b) and with

$$\omega_1(t) = \frac{\kappa^2}{|\phi|^2}\left\{\chi + \frac{\phi''}{2\phi} - \frac{3\phi'^2}{4\phi^2}\right\}.$$

Now when $|\rho|$ is very large the first of the two terms ρ^2 and $\omega_1(t)$ greatly overshadows the second and effectively minimizes its influence. Since in the complete absence of $\omega_1(t)$ the equation would be solved by the functions $e^{\pm\rho t}$, the suggestion is evidently compelling that the equation as it stands may be solvable, at least approximately, by an expression

$$V_n(t, \rho) \equiv e^{\rho t}S(t, \rho),$$

with S some polynomial in $1/\rho$, say

$$S(t, \rho) \equiv 1 + \frac{\alpha_1(t)}{\rho} + \frac{\alpha_2(t)}{\rho^2} + \cdots + \frac{\alpha_n(t)}{\rho^n}.$$

This expression is found directly to satisfy a differential equation

$$V_n'' - \left\{\rho^2 + \omega_1(t) + \frac{\Theta}{S}\right\} V_n = 0,$$

in which Θ is explicitly the sum of terms

$$\Theta \equiv (2\alpha_1' - \omega_1) + \sum_{j=2}^{n}\left(\frac{2\alpha_j' - \alpha_{j-1}\omega_1 + \alpha_{j-1}''}{\rho^{j-1}}\right) + \left(\frac{\alpha_n'' - \alpha_n\omega_1}{\rho^n}\right),$$

and this conclusion follows irrespective of what the coefficients $\alpha_j(t)$ may be. It is at once clear, however, that with such coefficients as fulfill the relations

$$(4) \qquad \alpha_1' = \frac{1}{2}\omega_1, \quad \alpha_j' = \frac{1}{2}(\alpha_{j-1}\omega_1 - \alpha_{j-1}''), \qquad (j = 2, 3, \cdots, n),$$

all terms of Θ but the last vanish. Now the successive determination of functions $\alpha_j(t)$ which do satisfy the relations (4) involves at each step at most a quadrature and is therefore a matter of no theoretical difficulty, at least if $\omega_1(t)$ is absolutely integrable over (t_a, t_b). This condition is surely fulfilled if the interval is finite. It is seen thus how a form V_n may be constructed to solve a differential equation

$$V_n'' - \left\{\rho^2 + \omega_1(t) + \frac{\alpha_n'' - \alpha_n\omega_1}{S\rho^n}\right\} V_n = 0,$$

and since this equation resembles the equation (3) and does so the more closely the larger n and $|\rho|$ are taken, the presumption becomes strong that, with $V_n(t, \rho)$ determined as indicated, the latter is an asymptotic representation of a true solution of the equation (3). Inasmuch as this latter equation depends only on ρ^2, the same might naturally be expected also of the form $V_n(t, -\rho)$. With proper adjustments these facts can indeed be shown.† I shall, however, omit a proof at this point, particu-

† J. Horn, *Ueber eine lineare Differentialgleichung zweiter Ordnung mit einem willkürlichen Parameter*, Mathematische Annalen, vol. 52 (1899), p. 271, and other papers by this author of about this date. G. D. Birkhoff, *On the asymptotic character of the solutions of certain linear differential equations containing a parameter*, Transactions of this Society, vol. 9 (1908), p. 219. O. Blumenthal, *Ueber asymptotische Integration linearer Differentialgleichungen*, etc., Archiv der Mathematik und Physik, vol. 19 (1912), p. 136.

larly as I shall later sketch out considerations in which such a proof is effectively included.

The facts to be noted are that under the conditions assumed a pair of solutions $v_\pm(t)$ of the equation (3) exist which fulfill a set of relations

$$|v_\pm(t) - V_n(t, \pm\rho)| < \frac{K}{\rho^{n+1}}|e^{\pm\rho t}|,$$

with K some constant and t ranging over the interval (t_a, t_b), or, to summarize in terms of the original variables:

If $\phi^2(x)$ is bounded from zero on (a, b) and the integral

$$\int_a^b \left| \frac{\chi}{\phi} + \frac{\phi''}{2\phi^2} - \frac{3\phi'^2}{4\phi^3} \right| dx \tag{5}$$

converges, then the equation (2) *has a pair of solutions $u_\pm(x)$, which over the interval (a, b) have the forms*

$$u_\pm(x) = \frac{1}{|\phi|^{1/2}} e^{\pm\lambda i\int_*^x \phi dz}\left\{1 + \frac{\beta_1(x)}{\kappa\lambda} + \frac{\beta_2(x)}{(\kappa\lambda)^2} + \cdots + \frac{\beta_n(x) + \epsilon_\pm(x, \lambda)}{(\kappa\lambda)^{n+1}}\right\}. \tag{6}$$

When $\phi^2 < 0$ on (a, b), the forms (6), with $n = 1$, show that there are solutions representable by the formulas

$$u_\pm(x) = \frac{1}{|\phi|^{1/2}} e^{\pm\lambda\int_*^x |\phi| dz}\{1 + O(1/(\kappa\lambda))\}. \tag{7a}$$

On the other hand, if $\phi^2 > 0$, a suitable combination of these forms yields the representation

$$u_\gamma(x) = \frac{1}{\phi^{1/2}}\left\{\cos\left(\lambda\int_*^x \phi dx + \gamma\right) + O(1/(\kappa\lambda))\right\}, \tag{7b}$$

in which γ may be any constant. The symbol $O(1/(\kappa\lambda))$ stands in each case for a function of x and λ which on (a, b) is uniformly of the order of $1/(\kappa\lambda)$ when $\kappa\lambda$ is sufficiently large.

I have supposed in this statement and throughout that λ has been taken real and positive. The formal work needs no modification when it is thought of as complex. The proof, how-

ever, requires that the quantity $\lambda i \int_a^x \phi dx$ remain in either a right or a left hand half of the complex plane. A representation (6) is, to be sure, obtained in either case, but the question of the relation between the solutions u which are involved in the two cases is left open.

I need perhaps hardly remark that once the forms (6) are known the change of variables used above need no longer be taken into account. The coefficients $\beta_j(x)$ may be successively determined by substitution of the forms directly into the equation (2).

It may at this point be appropriate for me to outline also a variation of the procedure above. You will, however, recognize it as of only formal and minor significance. If a solution of the equation (2) is assumed in the form

$$u(x) = e^{\lambda \int_a^x \sigma(x,\lambda)\, dx},$$

there results for σ the Riccatti equation

$$\frac{\sigma'}{\lambda} + \sigma^2 + \phi^2 - \frac{\chi}{\lambda^2} = 0,$$

and if σ is taken as a series in $1/\lambda$, that is,

$$\sigma(x, \lambda) = \sum_{j=0}^{\infty} \frac{\sigma_j(x)}{\lambda^j},$$

the differential equation becomes

$$(\sigma_0^2 + \phi^2) + \frac{1}{\lambda}(2\sigma_1\sigma_0 + \sigma_0') + \frac{1}{\lambda^2}(2\sigma_2\sigma_0 + \sigma_1^2 + \sigma_1' - \chi)$$
$$+ \sum_{j=3}^{\infty} \frac{1}{\lambda^j}\left(\sigma_{j-1}' + \sum_{n=0}^{j} \sigma_n \sigma_{n-j}\right) = 0.$$

By choosing the coefficients $\sigma_j(x)$ so as to make the individual terms zero, that is,

$$\sigma_0 = \pm i\phi, \qquad \sigma_2 = -\frac{1}{2\sigma_0}(\sigma_1' + \sigma_1^2 - \chi),$$
$$\sigma_1 = -\frac{1}{2\sigma_0}\sigma_0', \qquad \sigma_j = -\frac{1}{2\sigma_0}\left(\sigma_{j-1}' + \sum_{n=1}^{j-1} \sigma_n \sigma_{j-n}\right), \tag{8}$$

where $j=3, 4, \cdots$, the equation is formally solved. This form of procedure† is in some vogue among the physicists, to whom it is known as part of the so called "Wentzel, Kramers, Brillouin," (W.K.B.), method. I shall have occasion to refer to this further.

The formulas (6) each include a function which has been designated by ϵ, and which is not specified beyond the statement that it approaches zero as $\kappa\lambda$ becomes infinite. The presence of these terms represents, therefore, a certain degree of vagueness in the formulas. This is a general characteristic of asymptotic representations and is not to be neglected, for to do so may easily lead into an utter misuse of the formulas. In most actual cases more precise analysis may be made to yield specific bounds for the magnitudes of the ϵ terms. With such bounds at hand and with a definite margin of permissible error given, the condition that the vagueness fall within the allowed margin must be looked upon as fixing a lower bound, say N, to the values of $\kappa\lambda$ for which the formulas may be used. In any application of them it must then be regarded as a tacit hypothesis that $|\kappa\lambda| \geqq N$.

4. *Consideration of a Point where $\phi^2(x)$ is Zero.* Let us suppose now that at a point of the x axis, say at x_0, the function $\phi^2(x)$ becomes zero, and to fix the ideas let us suppose that the zero is a simple one and that ϕ^2 changes from negative to positive with increasing x. Then in any fixed closed interval on the left of but not including x_0 there are solutions represented by the forms (7a) when $\kappa\lambda$ is large. Every solution is therefore represented by some linear combination of these forms, with coefficients which are free from x, and so the solutions are to be characterized as of an *exponential* type. On the other hand, in any closed fixed interval on the right of and not including x_0, every solution is represented by a formula (7b) and so is of an *oscillatory* type. I wish for a moment to focus the attention upon such an interval, let us say for definiteness one on the right of x_0, with a view toward determining how large the interval (x_0, a) must be in order that the forms may be valid for $x \geqq a$.

The deductions of §3 were built upon the assumption that

† G. Wentzel, *Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik*, Zeitschrift für Physik, vol. 38 (1926), p. 518.

$|\rho|$ was sufficiently large to bring the vagueness of the resulting forms within the admissible bounds of error. The symbols $*$ and κ remained unspecified except that they were to be constant. Let us choose these values now as

$$* = x_0, \qquad \kappa = \int_{x_0}^{a} |\phi|\, dx.$$

The point t_a then remains fixed independently of the value a, the reasoning used in §3 applies, and the forms (6) are deducible. Thus the condition under which these forms suitably represent solutions is that $\lambda\int_{x_0}^{a}|\phi|\,dx$ be sufficiently large, and this clearly permits the point a to be taken closer and closer to x_0 as λ is taken larger and larger. Since the value of ξ, where

$$\xi = \lambda \int_{x_0}^{x} \phi\, dx, \tag{9}$$

increases numerically as x recedes in either direction from x_0, it follows that the forms (6) or (7) are usable in intervals in which $|\xi| \geqq N$. I have, of course, assumed here that the intervals are ones on which ϕ^2 is bounded from zero except near x_0.

With a value of N prescribed, the relation $|\xi| = N$ defines two points, say $x_{-N} < x_0$ and $x_N > x_0$. As we have seen, these points depend upon λ and approach x_0 as $1/\lambda$ approaches zero. The interval (x_{-N}, x_N) thus separates any one in which a solution is described by a combination of the forms (7a) from one in which it is of the form (7b), and hence in this interval the transition of a solution from the exponential to the oscillatory type takes place. That no information as to the mode of this transition is likely to be obtainable from the forms (7) themselves becomes immediately evident when it is observed that without exception these forms become infinite as x_0 is approached. The differential equation satisfied by their leading terms is found directly to be

$$U''(x) + \left\{\lambda^2\phi^2 + \frac{\phi''}{2\phi} - \frac{3\phi'^2}{4\phi^2}\right\} U = 0, \tag{10}$$

and this, unlike the equation (2), has a singularity at the point x_0. The interval (x_{-N}, x_N) is, therefore, evidently one in which no actual resemblance between the differential equations exists to be built upon.

It might still seem conceivable that by a passage through complex values encircling the point x_0 the representation valid for $x \geqq x_N$ might be found to serve also for $x \leqq x_{-N}$. Such expectations are shaken, however, when it is observed that if ϕ^2 is analytic the same is true of the solutions $u(x)$, whereas the forms (6) fail even to be single-valued in the region about x_0.

5. *The Stokes Phenomenon in the Theory of the Bessel Functions.* The differential equation for the Bessel functions $J_\nu(\xi)$, considered in the complex ξ plane, has two singular points located respectively at $\xi = 0$ and $\xi = \infty$. For large values of $|\xi|$ asymptotic forms for these functions have been deduced, generally through the medium of contour integral solutions of the equation, and when $\arg \xi$ is fixed, the functions are accordingly known to be represented by formulas

$$(11) \qquad J_{\pm\nu}(\xi) \sim \frac{1}{(2\pi\xi)^{1/2}} \left\{ C_{\pm 1} e^{i\xi} P\left(\frac{1}{\xi}\right) + C_{\pm 2} e^{-i\xi} P\left(\frac{-1}{\xi}\right) \right\},$$

where P denotes a formal power series with initial term 1, and the coefficients are appropriate constants.†

Now if the constants C are determined so that the formulas are correct, let us say for real positive values of ξ, then they will in general not be correct when ξ is negative. More generally speaking, if $\arg \xi$ is allowed to vary, say to increase steadily, then it is found that ever and anon the formulas theretofore valid cease to remain so, and that to retain them, the coefficients $C_{\pm 1}$ or $C_{\pm 2}$ must be changed. This is the Stokes phenomenon, and, as I have already indicated, its initial comprehension was not without difficulty. The functions on the left of formulas (11) are after all continuous in $\arg \xi$, whereas any change of a coefficient upon the right has every appearance of a discontinuous change in that member of the relation.

The explanation of the anomaly is nevertheless simple. To begin with, it may be observed that when $|\xi|$ is large the numerical value of $e^{i\xi}$ varies alternately between the large and the very small as $\arg \xi$ continues to increase. The dominance in the expression (11), therefore, passes alternately from the one term to the other. From here on I cannot do better than to quote directly from Stokes to the following effect: "The way in

† G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 1922, p. 202.

which the paradox · · · is explained is this. A semiconvergent series (considered numerically, and apart from its analytic form) defines a function only subject to a certain amount of vagueness which is so much the smaller as the modulus of the variable · · · is larger. · · · For critical values of θ [that is, $\arg\xi$] the vagueness of the superior function becomes sufficient to swallow up the inferior function. As θ passes through the critical value the inferior term enters as it were into a mist, is hidden for a little from view, and comes out with its coefficient changed."

It is evident at once that the first term in formulas (11) overshadows the second in the lower half-plane, and that the converse is true in the upper half-plane. We may think, therefore, of certain lines extending from the origin to ∞, one in each upper and lower half-plane, in crossing which the coefficients are to be changed. I shall refer to them as Stokes' lines. The precise location of such a line is never a matter of importance, since the term which is overshadowed remains so in a sector of which the axis of imaginaries is the bisector and which may be anything definitely less than π in magnitude.

Finally, I shall record for subsequent use two sets of constants which are appropriate to the formulas (11) when $\nu=1/3$, namely,

$$\text{(11a)}\quad \begin{aligned} &C_{\pm 1} = e^{\pi i(-1/4\mp 1/6)},\quad C_{\pm 2} = e^{\pi i(1/4\pm 1/6)},\quad (-\pi < \arg\xi < \pi);\\ &C_{\pm 1} = e^{\pi i(-1/4\mp 1/2)},\quad C_{\pm 2} = e^{\pi i(1/4\pm 1/6)},\quad (0 < \arg\xi < 2\pi). \end{aligned}$$

6. *A Particular Differential Equation.* Perhaps the simplest equation of the type (2) in which $\phi^2(x)$ has a zero, is the differential equation

$$y'' + \lambda^2 x y = 0. \tag{12}$$

It is explicitly solved by the functions

$$y_\pm(x) = x^{1/2} J_{\pm 1/3}(\xi), \qquad \text{with} \qquad \xi = \frac{2}{3}\lambda x^{3/2}. \tag{13}$$

Its study, because of this, involves a minimum of difficulties with a maximum of explicitness, which accounts largely for the fact that the equation has come to play a rather central role in the general theory. I shall devote the present section to a discussion of it.

If η is constant with respect to x, and is arbitrary except that it is restricted to the range $-\pi/2<\eta\leqq\pi/2$, the formula

$$y_\eta(x) = \Psi\lambda^{1/6}\left(\frac{2\pi}{3}\right)^{1/2}\xi^{1/2}\left\{\cos\left(\frac{\pi}{3}-\eta\right)J_{-1/3}(\xi) + \cos\left(\frac{\pi}{3}+\eta\right)J_{1/3}(\xi)\right\}, \tag{14}$$

in which $\Psi=(2/3)^{1/6}$, gives, except for a constant factor, the general solution of the differential equation. When x is real and positive the same is true of ξ; while negative values of x, say $\arg x=\pi$, yield values of ξ for which $\arg\xi=3\pi/2$, namely, $i\xi=|\xi|$. The first and second of the sets of coefficients (11a) are, therefore, to be associated respectively with the positive and negative values of x in the applications of the formulas (11). In the case at hand, that is, of equation (12), the point x_0 is the origin and the points at which $|\xi|=N$ are $x_{\pm N}=(3N/(2\lambda))^{2/3}$.

From the formulas (11), the asymptotic representations of $y_\eta(x)$ are found to be as follows: when $x\leqq x_{-N}$,

$$y_\eta(x) = \frac{1}{|x|^{1/4}}\left\{\sin\eta e^{|\xi|}\left\{1+O\left(\frac{1}{|\xi|}\right)\right\} + \frac{\cos\eta}{2}e^{-|\xi|}\left\{1+O\left(\frac{1}{|\xi|}\right)\right\}\right\}, \tag{14a}$$

and when $x\geqq x_N$,

$$y_\eta(x) = \frac{1}{x^{1/4}}\left\{\cos\left(\xi-\frac{\pi}{4}+\eta\right)+O\left(\frac{1}{\xi}\right)\right\}, \tag{14b}$$

and these involve a number of features which are worth noting. To begin with, it is clear from (14a) that as $x\to-\infty$, and hence $|\xi|\to+\infty$, the solutions with just one exception become infinite. The exceptional one is that for which $\eta=0$, that is, $y_0(x)$, and this approaches the limit zero. It is the only solution which is bounded for all values of x.

Like all asymptotic representations the formulas (14) involve an element of vagueness in the presence of the terms designated by $O(1/\xi)$. These terms are largest at the point x_{-N} or x_N, as the case may be, and if the formulas are to be taken as specific or

explicit to a high degree, and this is indeed a tacit assumption when only the leading terms are explicitly written down, then it may be inferred that N is large and the terms $O(1/\xi)$ so small whenever $|\xi| \geqq N$ that they are for most purposes negligible.

Let us consider the case in which $\eta \neq 0$ and is such that at some point x_{-T}, where $x_{-T} < x_{-N}$, the relation

$$\tan \eta = \frac{1}{2} e^{-2|\xi - T(x)|} \tag{15}$$

is fulfilled. Then at this point the explicit parts of the two terms in formula (14a) are equal. As x moves thence to the right, the second term rises rapidly to dominance and its margin of vagueness soon swallows up the first term entirely. As x moves to the left, on the other hand, the second term rapidly diminishes and is soon submerged in the margin of vagueness of the dominant first term. With the possible exception of the immediate neighborhood of the point x_{-T}, the graph of the solution accordingly *falls* when $x < x_{-T}$, and *rises* on the interval $x_{-T} < x \leqq x_{-N}$. The subscript T was chosen to suggest x_{-T} as such a point of transition.

An interval (x_{-T}, x_{-N}) evidently exists only if η is small enough to satisfy a relation (15), that is, if $\eta < \eta_0$, where

$$\tan \eta_0 = \frac{1}{2} e^{-2N}. \tag{15a}$$

If we agree to think of x_{-T} as coincident with x_{-N} when $\eta \geqq \eta_0$, and likewise to think of x_{-T} as $-\infty$ when $\eta = 0$, the description of the solutions given in (14a) and (14b) may be recast into the form

$$y_\eta(x) = \frac{\sin \eta}{|\phi|^{1/2}} e^{|\xi|}\{1 + O(1/|\xi|)\}, \qquad \text{when} \quad x < x_{-T}, \tag{16a}$$

$$y_\eta(x) = \frac{\cos \eta}{2|\phi|^{1/2}} e^{-|\xi|}\{1 + O(1/|\xi|)\}, \qquad \text{when} \quad x_{-T} < x \leqq x_{-N}, \tag{16b}$$

$$y_\eta(x) = \frac{\{1 + O(1/\xi^2)\}}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{4} + \eta + O(1/\xi)\right), \qquad \text{when} \quad x \geqq x_N. \tag{16c}$$

The formula (16c) is readily identified as equivalent to (14b). I have written $\phi^{1/2}$ in place of $x^{1/4}$, inasmuch as the equation (12) is of the type (2) with $\phi^2 = x$.

It is of some interest to note the inferences which can be drawn, especially with regard to the *phase constant* η, from the mere existence of representations identifiable with the ones thus given.

INFERENCE 1. If a solution $y(x)$ is of the form (16b) for all $x \leqq x_{-N}$, then $\eta = 0$.

INFERENCE 2. If a solution $y(x)$ is of the form (16b) in some interval adjacent to x_{-N}, then $\eta < \eta_0$.

INFERENCE 3. If a solution $y(x)$ is of the form (16a) for $x \leqq x_{-N}$, then $\eta \geqq \eta_0$.

INFERENCE 4. Since by taking x sufficiently large the margin of vagueness of formula (16c) may be reduced to any degree, a representation (16c) determines η and therefore the appropriate corresponding representations (16a) and (16b).

The differential equation

$$y'' + \lambda^2 x^\nu y = 0, \tag{17}$$

which is more general than (12), is likewise explicitly solvable, namely, by the functions

$$y_\pm = x^{1/2} J_{\pm 1/(\nu+2)}(\xi), \qquad \text{with} \qquad \xi = \frac{2}{\nu+2} \lambda x^{(\nu+2)/2}. \tag{17a}$$

A discussion of it proceeds in a manner similar to that of equation (12) which has been given, and may be made the basis of the deductions concerning the general equation (2) when $\phi^2(x)$ becomes zero to the degree ν. The index ν need not be an integer, but may be any real non-negative constant. In particular, the case $\nu = 0$ is that in which ϕ^2 is always of the same sign.

7. *The Method of Jeffreys and Kramers.* The first attack upon the equation (2) with x ranging over an interval upon which $\phi^2(x)$ has a zero was made, so far as I know, by Jeffreys in 1923.†

† H. Jeffreys, *On certain approximate solutions of linear differential equations of the second order*, Proceedings of the London Mathematical Society, (2), vol. 23 (1923), p. 428.

He was led to it in an investigation of the oscillations of surfaces of water, his particular problem leading to the Mathieu equation

$$u'' + (R - 2h^2 \cos 2x)u = 0.$$

When $|R| < 2h^2$, and x is on a suitable interval, this is evidently an equation of the type(2). I shall sketch out Jeffreys' procedure, which he restricts to the case in which the zero of $\phi^2(x)$ is simple and in which $\chi(x) \equiv 0$, that is, to the equation

$$u'' + \lambda^2 \phi^2(x) u = 0. \tag{2a}$$

There is no objection to taking the origin at the zero of ϕ^2, nor to supposing that the positive x direction is chosen so that ϕ^2 has always the same sign as x. Finally, the slope of $\phi^2(x)$ at the origin may be taken as unity since this merely fixes the specification of λ^2. It is to be made a hypothesis now that in a neighborhood of the origin the function $\phi^2(x)$ is represented with sufficient accuracy by a linear function. With the normalization agreed upon above this linear function will, of course, be x itself, and for small values of x we have $\xi = \lambda \int_0^x \phi(x)dx = (2/3)\lambda x^{3/2}$, approximately. At the same time the differential equation (2a) is approximated by the equation (12) of §6, namely,

$$y'' + \lambda^2 xy = 0. \tag{12}$$

Now when x is negative and sufficiently large the solutions of both the equations (2a) and (12) are asymptotically representable, as is shown by the formulas (7a) and (16). We are led, therefore, to write

$$u_-(x) \sim \frac{1}{|\phi|^{1/2}} e^{-|\xi|} \underset{\text{approx}}{=\!=} \frac{1}{|x|^{1/4}} e^{-(2/3)\lambda|x|^{3/2}} \sim 2y_0(x),$$
$$u_+(x) \sim \frac{1}{|\phi|^{1/2}} e^{|\xi|} \underset{\text{approx}}{=\!=} \frac{1}{|x|^{1/4}} e^{(2/3)\lambda|x|^{3/2}} \sim 2y_{\pi/6}(x). \tag{18a}$$

On the other hand, when the origin has been traversed and x is sufficiently removed therefrom on the positive side, the formulas (16c) and (7b) lead to

$$2y_0(x) \sim \frac{2}{x^{1/4}} \cos\left(\frac{2}{3}\lambda x^{3/2} - \frac{\pi}{4}\right)$$

$$\overline{\overline{\text{approx}}}\ \frac{2}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{4}\right) \sim 2u_{\pi/4}(x),$$

(18b)

$$2y_{\pi/6}(x) \sim \frac{2}{x^{1/4}} \cos\left(\frac{2}{3}\lambda x^{3/2} - \frac{\pi}{12}\right)$$

$$\overline{\overline{\text{approx}}}\ \frac{2}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{12}\right) \sim 2u_{\pi/12}(x).$$

The sequence of relations (18a) through (18b) is taken to supply the linear connections between the solutions of the formulas (7a) and (7b). These connections Jeffreys indicates in the form

(19)

$$\text{(a)} \quad \frac{1}{|\phi|^{1/2}} e^{-|\xi|} \longleftrightarrow \frac{2}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{4}\right),$$

$$\text{(b)} \quad \frac{1}{|\phi|^{1/2}} e^{|\xi|} \longleftrightarrow \frac{2}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{12}\right)\dagger,$$

where, to quote him: "the sign $\longleftrightarrow$ is used to indicate that the functions it connects are asymptotic approximations to the same function in different circumstances."

The method seems to me to be very simple and ingenious. There are, however, a number of points involved in it which will bear discussion. To begin with, the proper use of the symbol $\sim$ is inevitably accompanied by the tacit assumption that x is sufficiently removed from the origin, that is, to resort to our earlier notation, that $x \leqq x_{-N}$ or $x \geqq x_N$. The interval to be shunned, namely (x_{-N}, x_N), depends in length upon the magnitude of the permitted margin of vagueness in the representations, and is to be shortened only at the expense of precision. As against this, the proper use of the symbol $\overline{\overline{\text{approx}}}$ is subject to the contingency that $|x|$ is sufficiently small, let us say for definiteness that $x_{-k} \leqq x \leqq x_k$. The length of this interval is limited by the hypothesis that within it the difference between $\phi^2(x)$ and the linear approximating function be sufficiently small to keep

† Through an evident slip Jeffreys obtained on the right of this formula the incorrect form $(2/\phi^{1/2}) \cos(\xi - 5\pi/12)$.

under a permissible bound the cumulative error involved in the transfer from the solutions $u(x)$ to the solutions $y(x)$, and in the interchange between ϕ^2 and x, and between ξ and $(2/3)\lambda x^{3/2}$.

It is evidently a requirement of the method that the second of these intervals be large enough to contain the first within its interior, so that there may exist sub-intervals (x_{-k}, x_{-N}) and (x_N, x_k) in which the transfers may be made. It is not difficult to recognize that such an adjustment is always attainable by taking λ suitably large if no restriction upon the magnitude of that parameter exists.

Jeffreys appraises the method in the terms " . . . it has the advantage of rapid application in general and of practicability in many cases where no definite integral solutions (of the differential equation) can be found. On the other hand, it is not usually easy to fix rigorously the limits of error involved in adopting it." This seems to me to touch upon both the strength and the weakness of the procedure. As to the approximate magnitude of the error, Jeffreys concludes that it decreases with increasing λ and is of the order $O(\lambda^{-2/5})$. I must, however, confess myself unable to follow him to my own satisfaction in this final part of his deduction.

Concerning the form of the statement of results (19) I would object to the use of the symbol $\longleftrightarrow$, on the ground that it invites misconceptions which have not failed to show themselves in the literature. I find it not unnatural to read into a pair of forms connected by an arrow the thought that in the direction of the arrow the one form implies the other. In the case of the relations (19) this inference would certainly be incorrect. A review of the specific implications which were drawn from the analogous formulas (16), and which were listed in §6, will show the following, admitting that the formulas (19) ignore quantities of the order $O(1/\xi)$ and $O(1/\lambda)$.

Firstly, *the relation* (19a) *is one in which the left hand member implies and necessarily leads to the right hand member. The converse, however, is not true,* for it may be observed that in the deduction of the relation in question any solution $y_\eta(x)$ with η of a suitable smallness, may replace the solution $y_0(x)$ used, while such solutions are not all associated with the form of the left hand member given.

Secondly, *the relation* (19b) *is one in which the right hand*

member implies and necessarily leads to the left hand member, but not conversely.

Three years subsequent to Jeffreys, that is, in 1926, Kramers,† in a discussion of the Schrödinger wave equation for a particle with one degree of freedom, encountered and again discussed the equation of the type (2a). His discussion is in every essential feature identical with that of Jeffreys and results in the deduction of the formula (19a). A subsequent paper‡ adds to it the relation

$$\text{(19c)} \qquad \frac{1}{|\phi|^{1/2}}\, e^{|\xi|} \longleftrightarrow \frac{1}{\phi^{1/2}} \cos\left(\xi + \frac{\pi}{4}\right),$$

which is evidently the analog of the formula (16) in the particular case $\eta = \pi/2$. There is, of course, no contradiction between the relations (19b) and (19c), since, as I have just remarked regarding the implication of one form from the other, the left hand form may be inferred from the right hand one in each case, but not conversely. Among the physicists these deductions, and especially the formulas (19a) and (19c), are known as part of the "W.K.B." (Wentzel, Kramers, Brillouin) method, which I have already mentioned in an earlier connection in §3.

The extension of this method of procedure to the discussion of the differential equation of type (2a) in which $\phi^2(x)$ has a zero of any integral order was made by Goldstein§ in 1927. He relates his paper directly to that of Jeffreys and has similar applications in mind.

8. *The Method of Zwaan.*‖ A second method for dealing with the equation (2a) in the case that the zero of ϕ^2 is a simple one is built upon the idea of avoiding the critical point at which ϕ^2 vanishes, by encircling it in a passage through the complex plane. The procedure is the following, it being assumed by way

† H. A. Kramers, *Wellenmechanik und halbzählige Quantisierung*, Zeitschrift für Physik, vol. 39 (1926), p. 828.

‡ H. A. Kramers, and G. P. Ittmann, *Zur Quanteilung des asymmetrischen Kreisels* II, Zeitschrift für Physik, vol. 58 (1929), p. 217.

§ S. Goldstein, *A note on certain approximate solutions of linear differential equations of the second order*, etc., Proceedings of the London Mathematical Society, (2), vol. 28 (1928), p. 81.

‖ A. Zwaan, *Intensitäten im Ca-Funkenspektrum*, Dissertation, Utrecht, 1929.

of an explicit hypothesis, that within a suitable region of the complex plane about the zero of ϕ^2, the function $\phi^2(x)$ admits of approximation to the required degree of precision by means of a function which is analytic. For convenience I shall again suppose that the origin and the positive x direction have been chosen so that for real values of x the sign of the latter is always that of ϕ^2.

Let $u(x)$ be any solution of the differential equation (2a) which is real over the axis of reals. For negative values of x such that $x \leqq x_{-N}$ this solution is represented asymptotically by a linear combination of the forms (6), that is, with ξ defined as in formula (9),

$$u(x) \sim \frac{c_1}{\phi^{1/2}} e^{i\xi} + \frac{c_2}{\phi^{1/2}} e^{-i\xi}. \tag{20a}$$

For these negative values of x it will be recalled that $\arg \phi^2 = \pi$ and $\arg \xi = 3\pi/2$.

In the upper half of the complex x plane let a semi-circle be drawn with center at the origin and with a radius r, (i) such that $-r \leqq x_{-N}$ and $r \geqq x_N$, (ii) such that at each point of the arc an asymptotic representation (20a) exists, and (iii) such that within the semi-circle the hypothesis made concerning ϕ^2 is fulfilled. As x traverses this semi-circle in the clockwise sense the point ξ in the complex ξ plane traces out a curve Γ which begins on the negative axis of imaginaries, proceeds on the whole in the clockwise sense about the point $\xi = 0$, and ends on the positive axis of reals. We wish to consider the vicissitudes of the form (20a) during the transit of this arc.

Let it be supposed that the first term of the form (20a) is never dominant on the arc Γ. In this case the reality of the solution for the negative values of x requires that $\arg c_2 = \pi/4$. The continuity of the solution is not contrary to a suitably moderate change in the coefficient of the submerged term, say from c_1 to c_1^*, but would be irreconcilable with any change in c_2. Finally, the reality of the solution for positive values of x, where both ϕ^2 and ξ are positive, requires that in the end c_1^* and c_2 be complex conjugates. From this the ultimate form, valid for $x \geqq x_N$, is deduced to be

$$u(x) \sim \frac{|c_2|}{\phi^{1/2}} \left\{ e^{i\xi - i \arg c_2} + e^{-i\xi + i \arg c_2} \right\}. \tag{20b}$$

This, and the form (20a), give precisely the members of the relation (19a), which is thus established, in the sense that the left hand form implies the right. It will be noted that no deduction of the left hand form from the right by means such as outlined is possible, which accords with the earlier discussion of the relation in question.

The method was applied by Zwaan only when the coefficient c_1 in form (20a) is zero. It is at once evident that in most simple applications this yields an instance of the type considered, and hence affords a means, and a very simple one, for deriving the connection (19a).

The obvious advantage of the method lies in its great simplicity in cases in which it can be applied. Its weaknesses are several. Firstly, there is its excessively circumscribed range. Limited to the case of a simple zero of ϕ^2, there seems no possibility of obtaining from it any result whatever when the zero is of higher order. Secondly, there are involved in it many tacit assumptions, some by no means negligible, particularly as to the asymptotic character of the solutions in the complex plane. A critique and elaboration of this point is contained in the address recently delivered here at Chicago by Birkhoff.† Finally, an estimation of the relative error involved in the result is almost impossible, and unlike either the method of Jeffreys or that which I shall next discuss, the procedure yields no information regarding the functional character of a given solution in the neighborhood of a point where ϕ^2 becomes zero.

9. *Another Method.*‡ The procedures which I have outlined in the two preceding sections are evidently to be characterized as of the nature of means for knitting together the results given by the classical process of §3 for each of two separate portions, namely (a, x_{-N}) and (x_N, b), of the fundamental interval (a, b). Under them a solution of the differential equation is thought of as asymptotically *fitted* in the one sub-interval, likewise so

† G. D. Birkhoff, *Quantum mechanics and asymptotic series*, this Bulletin, vol. 39 (1933), p. 696.

‡ R. E. Langer, *On the asymptotic solutions of ordinary differential equations, with an application to the Bessel functions of large order*, Transactions of this Society, vol. 33 (1931), p. 23; and *On the asymptotic solutions of differential equations, with an application to the Bessel functions of large complex order*, Transactions of this Society, vol. 34 (1932), p. 447.

in the other, and devices are then added for patching the results together into a more or less complete representation. The method which I would now like to present to you attacks the problem from a more fundamental standpoint, that of obtaining at one and the same stroke a representation which is valid over the whole domain of the variable. Should there be no zero of the coefficient ϕ^2 in this domain, then the procedure of itself takes on the classical form. It is a method, therefore, which seeks to include and generalize rather than to extend by piecing out the method heretofore classical. Of its definiteness, adaptability, and scope I shall wish to say something anon.

A scrutiny of the special differential equation (17) and of its explicit solutions (17a) suggests the consideration in the general case of the functions

$$v_\pm \equiv \xi^\mu J_{\pm\mu}(\xi),$$

in which $J_{\pm\mu}$ stands, as customarily, for the Bessel functions of the indicated order, μ is a tentatively undetermined constant, and ξ is defined in the manner of formula (9). The differential equation satisfied by these functions is

$$\frac{d^2v}{d\xi^2} + \frac{(1-2\mu)}{\xi}\frac{dv}{d\xi} + v = 0,$$

or, in terms of x as the independent variable,

$$v'' + \left\{(1-2\mu)\frac{\lambda\phi}{\xi} - \frac{\phi'}{\phi}\right\}v' + \lambda^2\phi^2 v = 0.$$

If in this, finally, the term of the first order is removed, the substitution being

$$y = \Psi v, \qquad \text{with} \qquad \Psi = (\xi/\lambda)^{(1/2)-\mu}/\phi^{1/2},$$

the resulting equation is found to be

$$y'' + \left\{\lambda^2\phi^2(x) - \frac{\Psi''(x)}{\Psi(x)}\right\}y = 0, \tag{21}$$

of which the solutions are accordingly

$$y_\pm = \Psi(x)\xi^\mu J_{\pm\mu}(\xi). \tag{22}$$

Superficially the differential equation (21) is of the type of the equation (2) with which we are primarily concerned. That, however, was true also of the equation (10) which is solved by the leading terms of the classical asymptotic forms. The test lies in the functional character of the coefficient Ψ''/Ψ, which is given explicitly by the formula

$$-\frac{\Psi''}{\Psi}=\frac{\phi''}{2\phi}-\frac{3\phi'^2}{4\phi^2}+\left(\frac{1}{4}-\mu^2\right)\frac{\phi^2}{(\int_{x_0}^{x}\phi\,dx)^2}. \tag{23}$$

This evaluation shows that the difference between the equations (22) and (10) consists solely of the presence or absence of the term on the extreme right of (23). Near x_0, the zero of ϕ^2, the coefficient in equation (10) has a pole, a fact which presages the breakdown of the asymptotic representation as the value x_0 is approached. The salient fact now is that this pole, which for a random choice of the constant μ exists also in the function (23), may be obviated in the latter by the choice of μ as the number $1/(\nu+2)$, where ν is the order to which $\phi^2(x)$ becomes zero at x_0. This order may be any non-negative real number. The case $\nu=0$, with the consequent $\mu^2=1/4$, evidently causes the equation (21) to revert into identity with the equation (10), whereby the inclusion of the theory of §3 is at once made evident.

With the indicated choice of μ, the equation (21) yields an approximation to the given equation (2) over an entire interval including the critical point x_0 provided the coefficient ϕ^2 is otherwise bounded from zero. The nature of this approximation as it differs from that in the method of Jeffreys should be noted. There the difference between the two equations approaches zero, to be sure, at the critical point, but for every other fixed point it increases with λ to the order of λ^2. In the present case the difference between the equations is entirely independent of λ and is therefore the more effectively overshadowed by the term $\lambda^2\phi^2$ the larger the value of λ is taken.

It was remarked in §3 that the permissibility of infinite values for a or b, the limits of the interval, was contingent upon the convergence of the integral (5). The corresponding developments in the present method require the convergence of the integral

$$\int\left|\frac{\chi}{\phi}-\frac{\Psi''}{\phi\Psi}\right|dx,$$

over intervals not including the critical point x_0. Since the integrands differ only by the function

$$\left(\frac{1}{4} - \mu^2\right)\frac{\phi\, dx}{(\int_{x_0}^{x}\phi\, dx)^2},$$

which integrates directly into $(\mu^2-1/4)/\int_{x_0}^{x}\phi\, dx$, and since this approaches a limit as $|x|\to\infty$, it is clear that the generalized method requires none but the original condition in this respect.

The close resemblance between the differential equations (2) and (21), which becomes the more pronounced as λ increases, warrants the presumption that with suitable adjustments the solutions of the latter will represent asymptotically those of the former. I shall sketch out in the next section a mode of proof for this. In doing so I shall, however, content myself here with the consideration of the simplest case, namely with λ real and positive, with $\phi^2(x)$ real for real values of x, with the zero of ϕ^2 simple and located at the origin and with the sign of ϕ^2 the sign of x.

10. *Proof of the Representation.* The differential equation (2) may evidently be written if desired in the form

$$u'' + \left\{\lambda^2\phi^2 - \frac{\Psi''}{\Psi}\right\}u = \delta(x)\cdot u, \tag{2*}$$

with

$$\delta(x) = \chi(x) - \frac{\Psi''(x)}{\Psi(x)}.$$

For this equation, however, the part of the *reduced* equation is taken by (21), and on regarding (2*) formally as a *non-homogeneous* equation, a familiar procedure permits it to be *solved* in the form

$$u(x) = y(x) + \frac{1}{W}\int_{c}^{x} H(x, t)u(t)dt, \tag{24}$$

with

$$H(x, t) \equiv \{y_0(x)y_1(t) - y_1(x)y_0(t)\}\delta(t).$$

In this y_0 and y_1 may be any independent solutions of the equation (21); W is their Wronskian; c is any point of the interval

(a, b); and $u(x)$ is the solution of the equation (2) which *coincides* at $x=c$ with $y(x)$, any chosen solution of (21). I use the term *coincides* to mean that at the point in question the solutions take the same value, with derivatives which are also equal.

The equation (24) is an integral equation. If the entire right hand member is substituted for $u(t)$ in the integrand, and the process is repeated again and again, etc., there results formally the infinite development

$$u(x) = y(x) + \frac{1}{W}\int_c^x H(x, t)y(t)dt + \frac{1}{W^2}\int_c^x \int_c^s H(x, s)H(s, t)y(t)dt\,ds + \cdots . \tag{24a}$$

I will anticipate by mentioning that if $y(x)$ and c are suitably chosen, and λ is sufficiently large, this development converges and actually represents $u(x)$.

We are considering the case in which $\phi^2(x)$ has a simple zero, that is, in which $\mu=1/3$, and among these is included the special equation of §6. It will be found that the formula (14) gives the solutions of the equation (21), and without modifications other than such as are implicitly involved in the broader significance of the symbols. The asymptotic forms of these solutions are given by the formulas (16). With the special values $\eta=0$ and $\eta=\pi/2$, there are obtained thus a pair of solutions which may be used as y_0 and y_1 and which are respectively of the forms, when x is negative or positive, indicated by the relations:

$$y_0(x) \sim \frac{1}{2|\phi|^{1/2}} e^{-|\xi|} \sim \frac{1}{\phi^{1/2}} \cos\left(\xi - \frac{\pi}{4}\right),$$

$$y_1(x) \sim \frac{1}{|\phi|^{1/2}} e^{|\xi|} \sim \frac{1}{\phi^{1/2}} \cos\left(\xi + \frac{\pi}{4}\right).$$

The Wronskian of this pair has the value λ.

The first integral term on the right of the relation (24a) is explicitly the sum of the two expressions

$$I_1 = \frac{y_0(x)}{\lambda}\int_c^x \delta(t)y_1(t)y(t)dt,$$

$$I_2 = \frac{-y_1(x)}{\lambda}\int_c^x \delta(t)y_0(t)y(t)dt.$$

The points c and x have not as yet been specified. It may be observed, however, that if the range (c, x) contains wholly or partially the interval (x_{-N}, x_N), the integration over this latter contributes to the integral an amount which approaches zero with $1/\lambda$, for although the product of any two solutions $y(x)$ given by (14) has $\lambda^{1/3}$ as a factor, it is otherwise bounded and the interval itself is in length of the order of $\lambda^{-2/3}$. The integrations over the range where $|\xi(t)| \leqq N$ may, therefore, be disregarded henceforth. The considerations when $|\xi(t)| > N$ are with advantage shaped to differentiate the cases in which the solution $y(x)$ is *rising* at x_{-N} from those in which it is *falling* at that point.

Case 1. $y(x)$ *rising on* $c < x \leqq x_{-N}$. In this case, as was found in §6,

$$\tan\eta \leqq \frac{1}{2}e^{-2|\xi(c)|},$$

and from formulas (16)

$$(25a)\quad y(x) \sim \begin{cases} \dfrac{\cos\eta}{2|\phi|^{1/2}}e^{-|\xi|}, & \text{for } c < x \leqq x_{-N}, \\ \dfrac{1}{\phi^{1/2}}\cos\left(\xi - \dfrac{\pi}{4} + \eta\right), & \text{for } x \geqq x_N. \end{cases}$$

It is directly clear that when $|\xi(t)| > N$, the integrand in the expression I_1 is bounded, whence the same follows for the entire integral. The integrand in I_2, on the other hand, is of the form

$$\frac{\delta(t)}{\phi(t)}e^{-2i\xi(t)}O(1),$$

with $O(1)$ designating a bounded function. Since

$$|e^{-2i\xi(t)}| \leqq |e^{-2i\xi(x)}|, \qquad \text{whenever} \qquad t \leqq x,$$

it being immaterial whether t or x is positive or negative, the integral in question is seen to be of the form $e^{-2i\xi(x)}O(1)$. Thus

the sum I_1+I_2, that is, the second term on the right of (24a), is concluded to be of the form

$$\frac{O(1)}{\lambda^{5/6}}, \quad \text{when} \quad |\xi| \leqq N, \quad \text{and} \quad \frac{O(1)}{\lambda\phi^{1/2}} e^{-i\xi}, \quad \text{when} \quad |\xi| > N.$$

This mode of appraisement is extensible in a direct manner to the subsequent terms of (24a) and shows them to be of a similar type with successively higher powers of λ in the denominator, sufficient to cause convergence of the series. From this and the form of $y(x)$ it clearly follows that:

If $u(x)$, a solution of equation (2), *coincides with a rising solution $y(x)$ of equation* (21) *at a point c, where $c<x_{-N}$, then*

$$(26a)\quad u(x) = \begin{cases} \dfrac{\cos\eta}{2|\phi(x)|^{1/2}} e^{-|\xi|}\{1 + O(1/\xi) + O(1/\lambda)\}, & \\ \qquad\qquad\qquad\qquad \textit{when } c \leqq x \leqq x_{-N}, & \\ y(x) + \dfrac{O(1)}{\lambda^{5/6}}, \qquad \textit{when } x_{-N} \leqq x \leqq x_N, & \\ \dfrac{1}{\phi^{1/2}(x)}\left\{ \cos\left(\xi - \dfrac{\pi}{4} + \eta\right)\right. & \\ \qquad \left. + O(1/\xi) + O(1/\lambda)\right\}, \qquad \textit{when } x \geqq x_N, & \end{cases}$$

the phase constant η being that in the representation (25a).

Case 2. *$y(x)$ falling at x_{-N}.* In this case $\eta \geqq \eta_0$, with the latter given by (15a), and

$$(25b)\quad y(x) \sim \begin{cases} \dfrac{\sin\eta}{|\phi|^{1/2}} e^{|\xi|}, & \text{when } x \leqq x_{-N}, \\ \dfrac{1}{\phi^{1/2}} \cos\left(\xi - \dfrac{\pi}{4} + \eta\right), & \text{when } x \geqq x_N. \end{cases}$$

The integrands in the expressions I_1 and I_2 are respectively of the forms

$$\frac{\delta(t)}{\phi(t)} e^{2i\xi(t)} O(1), \qquad \text{and} \qquad O(1),$$

when $|\xi(t)| > N$, and if c is chosen as any point to the right of x_{-N}, the integrals are easily recognized as of the forms

$$e^{2i\xi(x)}O(1), \qquad \text{and} \qquad O(1).$$

The conclusion which follows as in Case 1 is the following:

If $y(x)$, a solution of equation (21), is falling at x_{-N}, and if $u(x)$, a solution of equation (2), coincides with $y(x)$ at a point c, where $x_{-N} \leqq c$, then

$$(26b)\quad u(x) = \begin{cases} \dfrac{\sin\eta}{|\phi(x)|^{1/2}}\, e^{|\xi|}\{1 + O(1/\xi) + O(1/\lambda)\}, & \text{when } x \leqq x_{-N}, \\ y(x) + \dfrac{O(1)}{\lambda^{5/6}}, & \text{when } x_{-N} \leqq x \leqq x_N, \\ \dfrac{1}{\phi^{1/2}(x)}\left\{\cos\left(\xi - \dfrac{\pi}{4} + \eta\right) + O(1/\xi) + O(1/\lambda)\right\}, & \text{when } x \geqq x_N, \end{cases}$$

with η determined by the relations (25b).

11. *Scope of the Method.* The proof which I have thus outlined applied especially to the case of a real parameter and variable and a simple zero of ϕ^2. I would emphasize, however, that these restrictions are in no way demanded by the method, but were imposed here for the sole purpose of avoiding tedious detail in the presentation. Generally the coefficient ϕ^2 may be any function, real or complex, which, in a given region R_x of the complex x plane, admits of representation in a form $(x-x_0)^\nu\phi_1^2(x)$, with ν any non-negative real number and $\phi_1^2(x)$ an analytic function which is bounded from zero. The parameter λ likewise need not be real. When such is the case the formula (9) defines ξ as a complex variable which ranges over a domain R_ξ obtainable by mapping from the original region. This domain is in general a portion of a Riemann surface with just one branch point located at $\xi=0$, with sheets in number finite or infinite depending upon the character of the number ν, and with boundaries depending upon λ and the boundaries of R_x.

The deduction of the approximating differential equation (21) is a formal matter unaffected by the greater or less generality of the elements it involves, and in the domain R_ξ, the solutions (22) are classically known for all values of ξ. As was remarked in §5, the asymptotic representations of these solutions are subject to the Stokes' phenomenon, which interposes itself in the transit of any upper or lower half-plane of the Riemann surface R_ξ. The sub-regions on this surface in which a single representation maintains, correspond to specific sub-regions of the original domain of x which are determinable from the relating map. Since, as in the simple case, the functions $y(x)$ represent asymptotically the properly associated solutions $u(x)$ of the differential equation under consideration, the quantitative details of the Stokes' phenomenon as it applies to the latter are revealed.

The procedure of which the foregoing is an outline has been carried out fully in the instance of several differential equations of classical importance. By suitable transformation, for example, the differential equation of the Bessel functions $J_\lambda(z)$ may be given the form of the equation (2) whenever $|\lambda|$ is large, and the structure of these important functions for all complex values of the variable and all large complex indices may therewith be derived.† The differential equation for the Hermite polynomials is a case of this kind, and a study of the solutions of this equation, or of its transform, the Weber equation, over the entire complex plane and for any large index, real or complex, has been successfully made.‡ Lastly, I have made a study by this method of the solutions of the Mathieu equation over the complex plane of the variable and with all relative configurations of the two (real) parameters with one of them large.§ In the instance of several other standard differential equations the method evidently makes a similar study possible.

12. *A More General Type of Differential Equation.* The function $\chi(x)$ in the coefficient of the equation (2) has throughout the discussion played a role of almost negligible importance. It

† R. E. Langer, loc. cit.

‡ N. Schwid, *The asymptotic forms of the Hermite and Weber functions*, Dissertation, Wisconsin, (1934).

§ R. E. Langer, *The solutions of the Mathieu equation with a complex variable and at least one parameter large*, Transactions of this Society, vol. 36 (1934).

will perhaps be granted in view of this that a dependence of this function upon λ is not significant so long as it remains bounded with respect to λ. It can, in fact, be shown that when the zero of ϕ^2 is a simple one the theory requires no essential modification even when χ is of the order of λ. That, however, is not so when the zero of ϕ^2 is of higher degree. Only the simplest of such cases, namely that of an equation

$$u'' - \{\lambda^2\phi^2(x) + \lambda\theta(x) + \chi(x, \lambda)\}u = 0, \tag{27}$$

in which ϕ^2 has a zero of the second order, has to my knowledge been at all investigated. Yet such equations have arisen in a number of connections and those studies which have been made have been made under the spur of impending applications.

Apparently the first discussion of an equation of the type (27) was given by Goldstein† in 1931. His method is an adaptation of that of Jeffreys discussed in §7. The variable and parameter are assumed to be real and the term $\chi(x, \lambda)$ does not occur in his equation. In consonance with the basic character of the method it is assumed that near x_0, the zero of ϕ^2, the functions ϕ^2 and θ are approximated to the requisite degree of accuracy by $c_1(x-x_0)^2$ and c_2, respectively, where c_1 and c_2 are constants. The approximating equation,

$$y'' - \{\lambda^2 c_1(x - x_0)^2 + \lambda c_2\}y = 0,$$

may be transformed into a Weber equation and with the solutions of this (which are of classically known forms) playing the role taken in §7 by the Bessel functions, the *connections* between the asymptotic representations which maintain on the two sides of x_0 are deduced.

The differential equation of this type has presented itself also in quantum mechanics, and in this connection I find a discussion of it to have been given by Voss in 1933.‡ There is apparently no essential difference between his treatment and that of Goldstein.

The method of §9 likewise lends itself to the generalization

† S. Goldstein, *A note on certain approximate solutions of linear differential equations of the second order*, (2), Proceedings of the London Mathematical Society, (2), vol. 33 (1932), p. 246.

‡ W. Voss, *Bedingungen für das Auftreten des Ramsaueffektes*, Zeitschrift für Physik, vol. 83 (1933), p. 581.

necessary for the discussion of the equation (27), and this, moreover, without restriction to real values of the variables.† Since the formulas as stated for the general case are of very considerable complexity, it might be worth noting here the forms which apply when the parameter and variable are real and the former is positive.

The equation is first to be adjusted so that the relation

$$3\phi'(x_0)\theta'(x_0) - 2\phi''(x_0)\theta(x_0) = 0$$

is satisfied. Then if x_0 is taken as the origin, and if ϕ^2 is bounded from zero except at the origin,‡ there are solutions describable as follows:

Case 1. $\phi^2(x) \geqq 0$: *solutions exponential.* With k and ξ defined by the relations

$$k = \frac{-\theta(0)}{4\phi'(0)}, \qquad \xi = \lambda \int_0^x \phi\, dx + \int_0^x \left\{ \frac{\theta}{2\phi} + \frac{k\phi}{\int_0^x \phi dx} \right\} dx,$$

there are solutions of the asymptotic forms

$$u_\alpha(x) \sim \begin{cases} \dfrac{(2\xi)^k e^{-\xi}}{\lambda^{1/4}\phi^{1/2}}, & \text{when } x \geqq x_N, \\ \dfrac{1}{\lambda^{1/4}\left|\phi\right|^{1/2}} \left\{ \dfrac{2\pi(2\xi)^{-k}e^{\xi}}{\Gamma(1/4-k)\Gamma(3/4-k)} - \delta\cdot(2\xi)^k e^{-\xi} \right\}, & \text{when } x \leqq x_{-N}, \end{cases}$$

$$u_\beta(x) \sim \begin{cases} \dfrac{1}{\lambda^{1/4}\phi^{1/2}} \left\{ \dfrac{2\pi(2\xi)^{-k}e^{\xi}}{\Gamma(1/4-k)\Gamma(3/4-k)} - \delta\cdot(2\xi)^k e^{-\xi} \right\}, & \text{when } x \geqq x_N, \\ \dfrac{1}{\lambda^{1/4}\left|\phi\right|^{1/2}} (2\xi)^k e^{-\xi}, & \text{when } x \leqq x_{-N}, \end{cases}$$

with $\delta = ie^{2k\pi i}$. It will be noted that $\delta = \pm 1$ when $k = n + 1/2 \pm 1/4$, that is, at the values of k for which the gamma functions become infinite. For such values a third solution which may be used is

† R. E. Langer, *The asymptotic solutions of certain linear ordinary differential equations of the second order*, Transactions of this Society, vol. 36 (1934), p. 90.

‡ For more precise statements I refer to Langer, loc. cit.

$$u_\gamma(x) \sim \begin{cases} \dfrac{1}{\lambda^{1/4}\phi^{1/2}}(2\xi)^{-k}e^{\xi}, & \text{when } x \geqq x_N, \\ \dfrac{\delta}{\lambda^{1/4}|\phi|^{1/2}}(2\xi)^{-k}e^{\xi}, & \text{when } x \leqq x_{-N}. \end{cases}$$

It is to be understood that arg ξ is to be taken as zero for both positive and negative values of x.

Case 2. $\phi^2(x) \leqq 0$: *solutions oscillatory.* With k and ξ given by the formulas

$$k = \frac{\theta(0)}{4|\phi'(0)|},$$

$$\xi = \lambda\int_0^x |\phi|\,dx - \int_0^x \left\{\frac{\theta}{2|\phi|} - \frac{k|\phi|}{\int_0^x |\phi|\,dx}\right\}dx,$$

there are solutions which when ξ is small are of the forms

$$u_\beta(x) = \frac{(2\xi)^{3/4}}{\lambda^{1/4}|\phi|^{1/2}}\left\{1 + \frac{4}{3}k\xi + \cdots\right\},$$

$$u_\epsilon(x) = \frac{(2\xi)^{1/4}}{\lambda^{1/4}|\phi|^{1/2}}\{1 + 4k\xi + \cdots\},$$

and which when ξ is large are of the forms

$$u_\beta(x) = \frac{(x/|x|)e^{k\pi/2}(\pi/|\phi|)^{1/2}}{\lambda^{1/4}|\Gamma(3/4 + ik)|}\cos\left\{\xi - k\log 2\xi + \arg\Gamma(3/4 + ik) - \frac{3\pi}{8}\right\},$$

$$u_\gamma(x) = \frac{2e^{k\pi/2}(\pi/|\phi|)^{1/2}}{\lambda^{1/4}|\Gamma(1/4 + ik)|}\cos\left\{\xi - k\log 2\xi + \arg\Gamma(1/4 + ik) - \frac{\pi}{8}\right\}.$$

The uncertainty or vagueness of these formulas is of the order of $1/\xi$ and of $(\log\lambda)/\lambda$.

13. *Remarks on the Wave Equation.* The Schrödinger wave equation for a particle with one degree of freedom is familiarly of the form

$$u'' + \frac{8\pi^2 m}{h^2}\{E - V(x)\}u = 0, \tag{28}$$

with m the mass of the particle, E and $V(x)$, respectively, the total energy and the potential energy, and h denoting Planck's constant. A wave function is a solution of the equation which remains bounded over the entire infinite axis of x. The cases of especial interest are those in which $E-V(x)$ changes sign for one or more values of x, and when such is the case the equation is evidently one of essentially the type we have been considering. In classical mechanics motion is possible only when $E-V(x)$ is positive. These are the intervals in which the solutions of the equation (28) are oscillatory. A wave function, that is, a permanently bounded solution, will in general exist only for special values of E, called characteristic values, the determination of which generally constitutes the essential part of the problem.

It is customary in the literature of the equation (28) to treat it as of the type of the equation (2) with the roles of λ^2 and ϕ^2 taken respectively by the expressions $4\pi^2/h^2$ and $2m\{E-V(x)\}$, and with $\chi \equiv 0$. The "W.K.B." procedure is then applied to supply the connection formulas across a zero of ϕ^2 for the representations of $u(x)$ asymptotic with respect to λ. In these representations the leading terms alone are generally considered. There are in this process a number of points to which I should like to call attention since, if nothing more, they hold the potentialities of error, and in the disregard of which conclusions have in a number of instances been drawn which seem to lack complete justification.

Firstly, the value assigned to λ, though in the prevailing units it may be large, is nevertheless fixed. The designation of a quantity as "of the order of $1/\lambda$" is therefore not a significant one, and formulas asymptotic with respect to λ should be used only with caution. It would seem that when such formulas are employed an actual verification that the relative vagueness of the formula is within the allowable margin of error is indispensable, and that this is particularly so when conclusions concerning quantities which may be relatively small are involved. Secondly, the function $2m\{E-V(x)\}$ is dependent upon the parameter E. In this respect it differs from the ϕ^2 of the general asymptotic theory. When the range of values of E is considera-

ble and also when E lies in certain critical ranges, this feature is easily conceived to be, and may be shown to be, one of first importance. Thirdly, the points at which the solutions change in character from the exponential to the oscillatory are not fixed but vary with the parameter E, a matter which is demonstrably not to be lightly ignored. For if, in particular, two such points approach each other too closely, the representations ordinarily obtained become entirely invalid.

Having raised these issues I would like to present a variation of the procedure outlined which seems to me to be more revealing, and logically much more defensible, though from the purely manipulative standpoint it is practically identical with the customary one. For convenience, I shall think of the case in which the function $2m\{E-V(x)\}$ is positive on a single finite interval, say $x_0 \leqq x \leqq x_0'$. The substitution

$$s = \frac{x - x_0}{x_0' - x_0}$$

gives to the equation (28) the form

$$\frac{d^2u}{ds^2} + \lambda^2\phi^2(E, s)u = 0, \tag{29}$$

in which we may choose

$$\lambda = \frac{2\pi}{h}\int_{x_0}^{x_0'} \{2m(E - V(x))\}^{1/2} \cdot dx,$$

$$\phi(E, s) \equiv \frac{\{E - V(s)\}^{1/2}}{\int_0^1 \{E - V(s)\}^{1/2} ds}. \tag{29a}$$

In this form almost all the questions raised above have been obviated. In the first place, λ is not fixed but increases with E. Secondly, though, to be sure, ϕ^2 still depends upon E, this dependence has been normalized, so to speak ironed out, to the extent that the area under the graph of ϕ between its zeros is fixed at the value 1. Thirdly, the zeros of ϕ^2 are *fixed*, specifically at $s=0$ and $s=1$. Evidently the equation in the form (29) differs in only a very minor way from the equation (2) of the general theory. Indeed, in the particular example $V(x) \equiv x^2$ which is often invoked, the equation is

$$\frac{d^2u}{ds^2}+\frac{128\pi^2mE^2}{h^2}(s-s^2)u=0,$$

which is precisely of the type considered.

From the form (29) the determination of the characteristic values proceeds in the usual way. The solution which is bounded on the range $-\infty<s<0$, must be that given by (26a) with $\eta=0$. Its form in a fixed interval entirely within (0, 1) must, therefore, be

$$\text{(30a)}\qquad \frac{1}{\phi^{1/2}(s)}\left\{\cos\left(\xi-\frac{\pi}{4}\right)+O(1/\lambda)\right\},$$

with $\xi=\lambda\int_0^s\phi\,ds$. Similar considerations, with the origin replaced by the point $s=1$, and the direction of the axis reversed, show that the solution bounded on $1<s<\infty$ must in the oscillatory interval have the form

$$\text{(30b)}\qquad \frac{\pm 1}{\phi^{1/2}(s)}\left\{\cos\left(\xi_1-\frac{\pi}{4}\right)+O(1/\lambda)\right\},$$

with $\xi_1=\lambda\int_s^1\phi\,ds$. Since $\xi+\xi_1=\lambda$, we may write (30b) in the form

$$\frac{1}{\phi^{1/2}(s)}\left\{\cos\left(\xi-\lambda-\frac{\pi}{4}\pm\frac{\pi}{2}\right)+O(1/\lambda)\right\},$$

and the condition that this be identical with (30a) is obviously that

$$\lambda=\left(n+\frac{1}{2}\right)\pi+O\left(\frac{1}{\lambda}\right).$$

Since λ must be large, this shows that the order of $1/\lambda$ is the order of $1/n$, and on substituting the value of λ the condition becomes

$$\text{(31)}\qquad \frac{2}{h}\int_{x_0}^{x_0'}\{2m(E-V(x))\}^{1/2}dx=n+\frac{1}{2}+O\left(\frac{1}{n}\right).$$

In order that a high degree of precision be assured, it is evidently requisite that n be sufficiently large.

If n is only moderate the result, at least without further considerations, is of little value. It may in this case be useful to vary

the method by *skipping over* the interval (x_0, x_0') entirely in the following way. Since the left member of (31) is not large, the value of θ given by the formula

$$\theta = \frac{2\pi}{h}(2m)^{1/2}\{V_m - E\},$$

with V_m denoting the minimum of $V(x)$, is of moderate magnitude. The equation (28) may, therefore, be written in the form

$$u'' - \{\lambda^2\phi^2(x) + \lambda\theta\}u = 0,$$

with

$$\lambda^2 = \frac{8\pi^2 m}{h^2}, \qquad \phi^2 = \{V(x) - V_m\},$$

and is thus reduced to one of the type referred to in §12.

The case in which the function $\{E - V(x)\}$ is positive on two or more distinct intervals has also important applications and has received considerable attention in the physical literature.† It would appear, however, that the considerations involved in it call for peculiar caution with respect to the limitations and implications of asymptotic representations, and these it seems have been all too often overlooked.

Let me suppose, to be concrete, that an equation (2) (or (29)) has a coefficient ϕ^2 which is negative on just one interval, say $x_1 < x < x_0$. On this interval the solutions are then of exponential type and in particular there are solutions of each of the forms

$$(32)\quad \text{(a)}\ \frac{1}{|\phi|^{1/2}}e^{|\xi_0|}, \qquad \text{(b)}\ \frac{1}{|\phi|^{1/2}}e^{-|\xi_0|}, \qquad \text{with} \qquad \xi_0 = \lambda\int_{x_0}^{x}\phi\,dx.$$

† See, for example, D. M. Dennison, and G. E. Uhlenbeck, *The two-minima problem and the ammonia molecule*, Physical Review, vol. 41 (1932), p. 313; H. D. Koenig, *Calculation of characteristic values for periodic potentials*, Physical Review, vol. 44 (1933), p. 657; Ta-You Wu, *Characteristic values of the two minima problem and quantum defects of states of heavy atoms*, Physical Review, vol. 44 (1933), p. 727.

It may not be concluded from this, however, that the solution represented by (32a) has the form

$$\frac{1}{\phi^{1/2}} \cos\left(\xi_0 + \frac{\pi}{4}\right), \tag{32a*}$$

for $x > x_0$, in virtue of the connection formula (19c), for that formula, as was remarked earlier, permits no implication from left to right. In fact, by (26b), the form beyond x_0 may be

$$\frac{\csc \eta}{\phi^{1/2}} \cos\left(\xi_0 - \frac{\pi}{4} + \eta\right),$$

with any η exceeding a small value η_0 (given by 15a). The form (31b) also implies strictly no more than the form

$$\frac{2 \sec \eta_1}{\phi^{1/2}} \cos\left(\xi_0 - \frac{\pi}{4} + \eta_1\right).$$

Since the value of η_1 in this is, however, restricted to be so small as to lie below the vagueness which has already been neglected in using only the leading term, the order of the error is not actually increased by suppressing η_1 entirely and writing

$$\frac{2}{\phi^{1/2}} \cos\left(\xi_0 - \frac{\pi}{4}\right), \tag{32b*}$$

which is in accord with the connection formula (19a). With this understanding the form (32b*) may be associated with any solution which is a *rising* exponential as it approaches the critical point. A situation similar to that at x_0 is, of course, also confronted in this matter at the point x_1.

The difficulty here seems to me to be circumventable only by *picking* special solutions in a suitable way, and holding fast to them throughout the discussion. I will try to make this clear as follows. Let $u_1(x)$ be chosen as that solution which for $x < x_1$ is of the form

$$\frac{1}{\phi^{1/2}} \cos\left(\lambda \int_x^{x_1} \phi \, dx + \frac{\pi}{4}\right).$$

It may be implied, for instance, by (26b) that for $x > x_1$ the form is

$$\frac{1}{|\phi|^{1/2}} e^{\lambda\int_{x_1}^{x}|\phi|dx}, \quad \text{that is,} \quad e^{\lambda\int_{x_1}^{x_0}|\phi|dx}\cdot\frac{1}{|\phi|^{1/2}} e^{-\lambda\int_{x}^{x_0}|\phi|dx}.$$

This solution as x_0 is approached from the left is *rising*. Hence its form for $x>x_0$ may be inferred, as in the case of (32b), and the complete relation indicated thus

$$\frac{1}{\phi^{1/2}}\cos\left(\lambda\int_{x}^{x_1}\phi\,dx+\frac{\pi}{4}\right)\underset{\text{at } x_1}{\longrightarrow}\frac{1}{|\phi|^{1/2}} e^{\lambda\int_{x_1}^{x}|\phi|dx}$$

$$\underset{\text{at } x_0}{\longrightarrow} e^{\lambda\int_{x_1}^{x_0}|\phi|dx}\cdot\frac{2}{\phi^{1/2}}\cos\left(\lambda\int_{x_0}^{x}\phi\,dx-\frac{\pi}{4}\right).$$

The inference has in each case been in the direction of the arrow.

For a second solution $u_2(x)$ we must now choose in the region $x>x_0$. In particular, the choice may be placed upon that one of the form

$$\frac{1}{\phi^{1/2}}\cos\left(\lambda\int_{x_0}^{x}\phi\,dx+\frac{\pi}{4}\right),$$

which then is certainly of the form

$$\frac{1}{|\phi|^{1/2}} e^{\lambda\int_{x}^{x_0}|\phi|dx}, \quad \text{that is,} \quad e^{\lambda\int_{x}^{x_0}|\phi|dx}\cdot\frac{1}{|\phi|^{1/2}} e^{-\lambda\int_{x_1}^{x}|\phi|dx},$$

when $x<x_0$. This, however, is *rising* as x approaches x_1 from the right, and hence for $x<x_1$ its form is as indicated in the formula

$$e^{\lambda\int_{x_1}^{x_0}|\phi|dx}\cdot\frac{2}{\phi^{1/2}}\cos\left(\lambda\int_{x}^{x_1}\phi\,dx-\frac{\pi}{4}\right)\underset{\text{at } x_1}{\longleftarrow}$$

$$\frac{1}{|\phi|^{1/2}} e^{\lambda\int_{x}^{x_0}|\phi|dx}\underset{\text{at } x_0}{\longleftarrow}\cdot\frac{1}{\phi^{1/2}}\cos\left(\lambda\int_{x_0}^{x}\phi\,dx+\frac{\pi}{4}\right),$$

the arrows again showing the direction in which the inference was made.

With a pair of solutions $u_1(x)$, $u_2(x)$ thus chosen, every solution is, of course, expressible as a combination

$$u(x) = Au_1(x) + Bu_2(x).$$

It is then given asymptotically in each of the three intervals by

the same combination of forms. In particular, in the interval $x>x_0$ we should have

$$u\sim\frac{1}{\phi^{1/2}}\left\{C\cos\left(\lambda\int_{x_0}^{x}\phi\,dx-\frac{\pi}{4}\right)+B\cos\left(\lambda\int_{x_0}\phi\,dx+\frac{\pi}{4}\right)\right\},$$

with

$$C=2Ae^{\lambda\int_{x_1}^{x_0}|\phi|dx},$$

or, as it may be alternatively written,

$$u\sim\left\{\frac{B^2+C^2}{\phi}\right\}^{1/2}\cos\left(\lambda\int_{x_0}^{x}\phi\,dx-\frac{\pi}{4}+\tan^{-1}\frac{B}{C}\right).$$

I have carried the formula to this last stage especially to warn against deductions from it which are not justified. For example, to conclude from it that the phase of u differs from that of u_1 by the amount $\tan^{-1}(B/C)$ would be warranted and meaningful only if this amount is large enough to project beyond the margin of vagueness which is inherent in all the formulas used. Though this is ignored when only the leading terms are written, it must nevertheless be borne in mind whenever conclusions are being drawn.

THE UNIVERSITY OF WISCONSIN

COMPOSITION OF ARITHMETICAL IDENTITIES

BY E. T. BELL

1. *Examples of Composition.* This concerns a new type of arithmetical identity involving parity functions. The letters n, d, δ, m, t, τ, with or without accents or suffixes, denote integers greater than zero; the n, d, δ are arbitrary, the m, τ, odd. Hence, if $n=t\tau$, t is a divisor of n whose conjugate divisor (τ) is odd. In what immediately follows, m, n (without suffixes or accents) are constant; summations refer to all values of the variable integers in the indicated partitions consistent with the notation just explained.

Let $f(x, y)$ be single-valued and finite for all pairs of integer values of x, y, and let $f(x, y)$ be even in each of x, y,

$$f(x, y) = f(-x, y) = f(x, -y).$$

Then there are the following known arithmetical identities, proofs of which will be indicated in §2.

(I) $\quad 2n = m_1 + m_2; \quad n = t\tau, \quad m_1 = t_1\tau_1, \quad m_2 = t_2\tau_2$:

$$\sum[f(t_1 + t_2, \tau_1 - \tau_2) - f(t_1 - t_2, \tau_1 + \tau_2)] = \sum t[f(2t, 0) - f(0, 2t)].$$

(II) $\quad n = n_1 + n_2; \quad n = d\delta, \quad n_1 = d_1\delta_1, \quad n_2 = d_2\delta_2$:

$$\sum[f(d_1 + d_2, \delta_1 - \delta_2) - f(d_1 - d_2, \delta_1 + \delta_2)]$$
$$= \sum\Big[(d-1)\{f(d, 0) - f(0, d)\} + 2\sum_{r=1}^{\delta-1}\{f(d, r) - f(r, d)\}\Big].$$

A sum in which the upper limit exceeds the lower is vacuous, by convention.

Repeating (I) in accented letters, we have (n' is constant)

(I′) $\quad 2n' = m_1' + m_2'; \quad n' = t'\tau', \quad m_1' = t'\tau_1', \quad m_2' = t_2'\tau_2'$:

$$\sum[f(t_1' + t_2', \tau_1' - \tau_2') - f(t_1' - t_2', \tau_1' + \tau_2')]$$
$$= \sum t'[f(2t', 0) - f(0, 2t')].$$

From (I) and (I′), and hence from (I) alone, we can infer by elementary arithmetical manipulations the following identity, which we number (I, I′) for a reason appearing in §2; n, n' are

constant; $\sigma_1(n') \equiv \zeta_1(n') + \zeta_1'(n')$, $\zeta_1(n') \equiv$ the sum of the divisors of n', $\zeta_1'(n') \equiv$ the sum of the odd divisors of n'.

$$
\begin{aligned}
\text{(I, I')}\quad & 2n = m_1 + m_2; \quad n = t\tau, \quad m_1 = t_1\tau_1, \quad m_2 = t_2\tau_2; \\
& 2n' = m_1' + m_2'; \quad n' = t'\tau', \quad m_1' = t_1'\tau_1', \quad m_2' = t_2'\tau_2' : \\
& \sum [f((t_1 + t_2)(t_1' - t_2'), (\tau_1 - \tau_2)(\tau_1' + \tau_2')) \\
& \quad - f((t_1 + t_2)(t_1' + t_2'), (\tau_1 - \tau_2)(\tau_1' - \tau_2')) \\
& \quad - f((t_1 - t_2)(t_1' - t_2'), (\tau_1 + \tau_2)(\tau_1' + \tau_2')) \\
& \quad + f((t_1 - t_2)(t_1' + t_2'), (\tau_1 + \tau_2)(\tau_1' - \tau_2'))] \\
& = \sum t[\sigma_1(n')f(0, 0) - t'\{f(4tt', 0) + f(0, 4tt')\}].
\end{aligned}
$$

In the same way (II), (I) give the next, in which n, n' are constant.

$$
\begin{aligned}
\text{(II, I')}\quad & n = n_1 + n_2; \quad n = d\delta, \quad n_1 = d_1\delta_1, \quad n_2 = d_2\delta_2; \\
& 2n' = m_1' + m_2'; \quad n' = t'\tau', \quad m_1' = t_1'\tau_1', \quad m_2' = t_2'\tau_2' : \\
& \sum [f((d_1 + d_2)(t_1' - t_2'), (\delta_1 - \delta_2)(\tau_1' + \tau_2')) \\
& \quad - f((d_1 + d_2)(t_1' + t_2'), (\delta_1 - \delta_2)(\tau_1' - \tau_2')) \\
& \quad - f((d_1 - d_2)(t_1' - t_2'), (\delta_1 + \delta_2)(\tau_1' - \tau_2')) \\
& \quad + f((d_1 - d_2)(t_1' + t_2'), (\delta_1 + \delta_2)(\tau_1' - \tau_2'))] \\
& = \sum t'[2(d - 1)f(0, 0) - (d + 2\delta - 3)\{f(2dt', 0) + f(0, 2dt')\} \\
& \quad + 2\sum_{r=1}^{\delta-1}\{f(2rt', 0) + f(0, 2rt')\}].
\end{aligned}
$$

As indicated by the notation, the t_1, τ_1 occurring in the arguments of f in (I) are conjugate divisors of m_1, and likewise for t_2, τ_2, and m_2. The point to be noticed is that the t_1, τ_1 occurring in the argument of f are conjugate, and are not taken from different decompositions of m_1. Similar remarks apply to all the identities.

The arguments in (I, I') are bilinear in the arguments occurring in (I) and (I'), and similarly for (II, I'). We call (I, I') the *composition* of (I) with itself, or the *duplication* of (I), and (II, I') the composition of (II) with (I). Note that order is relevant in the symbol (II, I'); it will be seen that (I', II) is meaningless. The general problem suggested by these examples can now be

stated. *Given s arithmetical identities concerning parity functions summed over quadratic partitions, to construct from them identities concerning parity functions whose arguments are of degree s, are products of the arguments of the functions in the given identities, and are linear in the arguments of the respective identities.* The resulting identities will be called *compositions* of the given identities. This may be extended in obvious ways to the compounding of compositions. It will be sufficient here to indicate the algebraic method by which the above examples were obtained. Although elementary arithmetical proofs can be given, the general algebraic method is preferable on account of its flexibility, suggestiveness in finding compositions, and ease of application.

2. *Algebraic Method.* The identity (I) is the paraphrase of the following identity* in elliptic theta functions:

$$\phi_{100}(x, y)\phi_{100}(x, -y) = \psi_{10}^2(x) - \psi_{10}^2(y),$$

$$\phi_{100}(x, y) \equiv \vartheta_1'\vartheta_1(x + y)/(\vartheta_0(x)\vartheta_0(y)), \tag{1}$$

$$\psi_{10}(x) = \vartheta_2\vartheta_3\vartheta_1(x)/\vartheta_0(x),$$

for which the requisite expansions are

$$\phi_{100}(x, y) = 4\sum_m q^{m/2}\sum \sin(tx + \tau y), \qquad (m = t\tau),$$

$$\psi_{10}^2(x) = 4\sum_n q^n[\sigma_1(n) - 2\sum t\cos 2tx], \quad (n = t\tau),$$

the m, n, t, τ notation being as in §1.

The identity (II) is the paraphrase of

$$\phi_{111}(x, y)\phi_{111}(x, -y) = \psi_{01}^2(x) - \psi_{01}^2(y), \tag{2}$$

$$\phi_{111}(x, y) \equiv \vartheta_1'\vartheta_1(x + y)/\vartheta_1(x)\vartheta_1(y), \psi_{01}(x) \equiv \vartheta_2\vartheta_3\vartheta_0(x)/\vartheta_1(x),$$

$$\phi_{111}(x, y) = \operatorname{ctn} x + \operatorname{ctn} y + 4\sum_n q^{2n}\sum \sin 2(dx + \delta y),$$

$$(n = d\delta),$$

$$\psi_{01}^2(x) = \csc^2 x + 4\sum_n q^n\sigma_1(n) - 8\sum_n q^{2n}\sum d\cos 2dx,$$

$$(n = d\delta).$$

* Several examples of the process of paraphrasing are given in the series of notes, *Quadratic partitions*, appearing in this Bulletin; the last note is in vol. 39 (1933). The series is to be continued.

Obviously

$$F(x, y) \equiv \phi_{100}(xu, yv)\phi_{100}(xu, -yv),$$

where u, v are independent variables, is finite and single-valued for all pairs of integer values of x, y, and

$$F(x, y) = F(-x, y) = F(x, -y).$$

Hence we may take $f(x, y) = F(x, y)$ in the left of (I) and, by (1),

$$f(x, y) = \psi_{10}^2(xu) - \psi_{10}^2(yv)$$

in the right of (I). Equating coefficients of like powers of q in the resulting identity, we obtain a trigonometric identity in u, v which paraphrases into (I, I′). In the same way (II, I′) follows from (2). Note that (I, II′) which, if it existed, would be the result of taking

$$f(x, y) = \phi_{111}(xu, yv)\phi_{111}(xu, -yv)$$

in the left of (I), and

$$f(x, y) = \psi_{01}^2(xu) - \psi_{01}^2(yv)$$

in the right, is meaningless, since $\phi_{111}(x, y)$ is undefined if $xy = 0$. From these examples the general process is evident. The symbolic algebra for performing the compositions, which is suggested by comparing the left hand members of the given identities with their compositions, will be developed in another paper, where further examples will be given. The partitions in a given set of identities to be compounded may be of any degrees (the degree is the number of independent variables). The number of independent variables may be increased or diminished by composition.

California Institute of Technology

CONTINUED FRACTIONS AND CROSS-RATIO GROUPS OF CREMONA TRANSFORMATIONS*

BY H. S. WALL

1. *Introduction.* An arbitrary cross ratio

$$r_{ijkl} = [z_i, z_j, z_k, z_l]$$

of four of the independent variables $z_1, z_2, \cdots, z_n$ is expressible rationally in terms of the ratios of any fundamental system such as

$$C: \qquad s_i = [z_{i-1}, z_{i-2}, z_i, z_{i-3}], \qquad (i = 4, 5, \cdots, n),$$

or

$$M: \qquad r_i = [z_1, z_2, z_3, z_i], \qquad (i = 4, 5, \cdots, n),$$

of $n-3$ independent ratios. If any particular system of $n-3$ independent ratios be associated with a particular order of the variables by varying the order of the n variables, we shall have in all $n!$ conjugate systems; these systems are expressible rationally in terms of the original system and in terms of any system of the set. Hence arises a group of $n!$ Cremona transformations on $n-3$ variables. E. H. Moore,† H. E. Slaught‡ and others have studied the group based on the initial system M.

In this paper I have shown that the transformations based on the system C have application to continued fractions of the form

$$\xi = 1 - \cfrac{x_1}{1 - \cfrac{x_2}{1 - \cfrac{x_3}{1 - \cdots}}}.$$

If we take for the n variables $z_1, z_2, \cdots, z_n$ any n consecutive convergents of the continued fraction, say

$$(1) \qquad z_i = \frac{A_{q+i}}{B_{q+i}}, \qquad (i = 1, 2, \cdots, n; q \geqq -1),$$

* Presented to the Society, under a somewhat different title, April 6, 1934.

† E. H. Moore, American Journal of Mathematics, vol. 22 (1900), pp. 279–291.

‡ H. E. Slaught, ibid., pp. 343–380, and Part II in vol. 23.

then in this case the fundamental system C is $n-3$ consecutive elements x_i of the continued fraction ξ, namely,

$$s_i = x_{q+i}, \qquad (i = 4, 5, \cdots, n). \tag{2}$$

The Cremona transformations are on these elements, and their effect is to permute n consecutive convergents of ξ among themselves. The transformations do not disturb the convergence or divergence of the continued fraction, and result therefore in new convergence criteria.

2. *The Group C_{n1} of Cremona Transformations.* Let $A_{m,t}/B_{m,t}$, $(t \geqq 0,\ A_{m,0}/B_{m,0} = A_m/B_m)$, be the mth convergent of the continued fraction

$$1 - \cfrac{x_{1+t}}{1 - \cfrac{x_{2+t}}{1 - \cfrac{x_{3+t}}{1 - \cdots}}}.$$

Then with the aid of the identity*

$$A_{n+t-1}B_{t-1} - A_{t-1}B_{n+t-1} = -x_1x_2 \cdots x_tB_{n-1,t},$$

we find that a cross ratio $r_{ijkl} = [z_i, z_j, z_k, z_l]$, $(i<j<k<l)$, of four of the n variables (1) can be expressed in the form

$$r_{ijkl} = \frac{B_{k-i-1,q+i+1}B_{l-j-1,q+j+1}}{B_{k-j-1,q+j+1}B_{l-i-1,q+i+1}}. \tag{3}$$

The other five distinct ratios of these four variables can be obtained from the ratio (3) by the well known transformations

$$r_{jikl} = \frac{1}{\lambda}, \qquad r_{ikjl} = 1 - \lambda, \qquad r_{kijl} = \frac{1}{1-\lambda},$$
$$r_{kjil} = \frac{\lambda}{\lambda - 1}, \qquad r_{jkil} = \frac{\lambda - 1}{\lambda}, \tag{3a}$$

where $\lambda = r_{ijkl}$. In particular, we find that

$$s_i = x_{q+i}, \qquad (i = 4, 5, \cdots, n). \tag{4}$$

Inasmuch as (3) depends upon $x_{q+4}, x_{q+5}, \cdots, x_{q+n}$ only, it follows that (4) is a fundamental system of cross ratios.

Let $j_1, j_2, \cdots, j_n$ be an arbitrary permutation of $1, 2, \cdots, n$; and put

* Perron, *Die Lehre von den Kettenbrüchen*, 1st ed., p. 17.

$$a = \begin{pmatrix} 1, & 2, \cdots, & n \\ j_1, & i_2, \cdots, & i_n \end{pmatrix}.$$

Then

$$C^a: \qquad s_i^a = [z_{j_{i-1}}, z_{j_{i-2}}, z_{j_i}, z_{j_{i-3}}], \qquad (i = 4, 5, \cdots, n),$$

is also a fundamental system; and C^a is expressible rationally in terms of C:

$$s_i^a = f_i^a(s_4, s_5, \cdots, s_n), \qquad (i = 4, 5, \cdots, n),$$

or, simply, $C^a = f^a C$.

If b is another permutation, then $C^{ab} = f^a C^b = f^{ab} C$. The $n!$ Cremona transformations $f^a, f^b, \cdots$ form a group $C_{n!}$ simply isomorphic with the symmetric permutation group $C_{n!}$.

The ratios s_i are given in terms of the ratios r_i of E. H. Moore by the formula

$$(5) \qquad r_{ijkl} = [r_i, r_j, r_k, r_l], \qquad (r_1 = \infty, r_2 = 0, r_3 = 1).$$

The r_i can be expressed in terms of the s_i by means of (3) and (3a). Thus the group $C_{n!}$ is equivalent to that of E. H. Moore under the transformation

$$s_i = [r_{i-1}, r_{i-2}, r_i, r_{i-3}], \qquad (i = 4, 5, \cdots, n).$$

3. *Convergence Criteria for Continued Fractions.* Let $C_{m!}'$ be the subgroup of $C_{2m!}$ corresponding to permutations of the form

$$a = \begin{pmatrix} 1 & 2 & \cdots & m & 1+m & 2+m & \cdots & m+m \\ j_1 & j_2 & \cdots & j_m & j_1+m & j_2+m & \cdots & j_m+m \end{pmatrix},$$

in which $j_1, j_2, \cdots, j_m$ is a permutation of $1, 2, \cdots, m$. It is plain that $f_4^a, f_5^a, \cdots, f_m^a$ are functions of $s_4, s_5, \cdots, s_m$ alone, while $f_{m+4}^a, f_{m+5}^a, \cdots, f_{2m}^a$ are functions of $s_{m+4}, s_{m+5}, \cdots, s_{2m}$ alone. The remaining three functions depend in general upon $s_4, s_5, \cdots, s_{2m}$.

Let $k \geqq 0$, $m \geqq 3$, $z_{i+1} = A_{km+i}/B_{km+i}$, $(i = 0, 1, \cdots, 2m-1)$. Then $s_i = x_{km+i-1}$, $(i = 4, 5, \cdots, 2m)$. We have

$$x_{km+i-1}^a = f_i^a(x_{km+3}, x_{km+4}, \cdots, x_{(k+2)m-1}), \qquad (i = 4, 5, \cdots, 2m),$$

with the exceptional values*

* Perron, loc. cit., p. 198.

$$x_0^a = \frac{A_{j_1-1}}{B_{j_1-1}}, \qquad x_1^a = \frac{A_{j_1-1}}{B_{j_1-1}} - \frac{A_{j_2-1}}{B_{j_2-1}},$$

$$x_2^a = \left(\frac{A_{j_2-1}}{B_{j_2-1}} - \frac{A_{j_3-1}}{B_{j_3-1}}\right) \Big/ \left(\frac{A_{j_1-1}}{B_{j_1-1}} - \frac{A_{j_2-1}}{B_{j_2-1}}\right).$$

Then the continued fraction (with real or complex elements)

$$\xi_a = x_0^a - \cfrac{x_1^a}{1 - \cfrac{x_2^a}{1 - \cfrac{x_3^a}{1 - \cdots}}}$$

has the same convergents as ξ, but in a different order. Consequently, if W_a is a region in m-space such that when the points

$$(x_{km}, x_{km+1}, \cdots, x_{(k+1)m-1}), \qquad (k = 0, 1, 2, \cdots),$$

range over W_a we shall always have*

$$|x_n^a| \leqq \frac{1}{4}, \qquad (n = 2, 3, 4, \cdots),$$

then ξ converges. In this manner every transformation of the group $C_{m!}$ gives a convergence theorem for ξ.

4. *The Groups* C_{120} and C_{120}'. As an illustration,† we shall consider the groups $C_{5!}$ and $C_{5!}'$. The group C_{120} is generated by the four transformations

$$K \sim (34), \qquad L \sim (23)(45), \qquad M \sim (45), \qquad T \sim (12).$$

The three transformations K, L, M generate by themselves a subgroup C_{24} of the main group C_{120}; T will extend C_{24} to the main group.

By (3) and (3a) the transformations K, L, M, T are found to be as follows:

$$K\colon x' = \frac{x}{x-1}, \quad y' = \frac{1}{y}; \qquad L\colon x' = \frac{1-y}{x}, \quad y' = y;$$

(6)

$$M\colon x' = \frac{x}{1-y}, \quad y' = \frac{y}{y-1}; \qquad T\colon x' = \frac{x}{x-1}, \quad y' = \frac{y}{1-x},$$

where, to avoid subscripts, we have put $s_4 = x$, $s_5 = y$.

---

* Perron, loc. cit., p. 259.

† The details of this illustration were worked out by Miss Lozelle Thomas.

In the xy-plane the geometrical configuration for the group C_{24} is as follows. The curves

$$1.\ x^2=1-y, \qquad 2.\ y=0, \qquad 3.\ x=0,$$

$$4.\ y=(x-1)^2, \qquad 5.\ x=1, \qquad 6.\ y=1,$$

$$7.\ x+y=1, \qquad 8.\ x^2+2xy-2x-y+1=0,$$

divide the plane into 24 regions. Take the fundamental region I to be the region in the second quadrant bounded by the curves 1, 2 and 3. The other regions in the second quadrant are then: L bounded by 1, 2, 6; LK bounded by 6, 8; LKL bounded by 7, 8; MKL bounded by 4, 7; and MK bounded by 3, 4.

In the first quadrant the regions are: K bounded by 3, 8; KL bounded by 5, 6, 8; $KMKL$ bounded by 4, 5, 6, extending to infinity; MLK bounded by 4, 6; $MLKM$ bounded by 2, 4, 6; $MLKLM$ bounded by 4, 5, 6; KLK bounded by 1, 5, 6; $LKLK$ bounded by 1, 7; $MKLM$ bounded by 4, 7; and KMK bounded by 2, 3, 4.

In the third quadrant the regions are: LM bounded by 1, 2; and M bounded by 1, 2, 3.

In the fourth quadrant the regions are: KM bounded by 2, 3, 8; KLM bounded by 5, 8; $MLKLK$ bounded by 1, 5; $MKLK$ bounded by 1, 7; $LKLM$ bounded by 7, 8; and LKM bounded by 2, 8.

The generators of C_{120}' correspond to the permutations (34)(89), (23)(45)(78)(9, 10), (45)(9, 10), and (12)(67). If we denote them by K', L', M', T', respectively, then these transformations are found to be

$$K': \quad x'=\frac{x}{x-1}, \quad y'=\frac{1}{y}, \quad \bar{x}'=\frac{\bar{x}}{\bar{x}-1}, \quad \bar{y}'=\frac{1}{\bar{y}},$$

$$u'=\frac{u}{u-1}, \quad v'=\frac{v}{1-w}, \quad w'=\frac{w}{1-\bar{x}};$$

$$L': \quad x'=\frac{1-y}{x}, \quad y'=y, \quad \bar{x}'=\frac{1-\bar{y}}{\bar{x}}, \quad \bar{y}'=\bar{y},$$

$$u'=\frac{1-y}{u}, \quad v'=\frac{v}{v+w-1}, \quad w'=\frac{w}{v+w-1};$$

$$M': \quad x' = \frac{x}{1-y}, \quad y' = \frac{y}{y-1}, \quad \bar{x}' = \frac{\bar{x}}{1-\bar{y}}, \quad \bar{y}' = \frac{\bar{y}}{\bar{y}-1},$$
$$u' = \frac{1}{u}, \quad v' = \frac{v}{v-1}, \quad w' = \frac{w}{1-v};$$

$$T': \quad x' = \frac{x}{x-1}, \quad y' = \frac{y}{1-x}, \quad \bar{x}' = \frac{\bar{x}}{\bar{x}-1}, \quad \bar{y}' = \frac{\bar{y}}{1-\bar{x}},$$
$$u' = \frac{u}{1-v}, \quad v' = \frac{v}{v-1}, \quad w' = \frac{1}{w},$$

where for simplicity we have put $x = x_{5n+3}$, $y = x_{5n+4}$, $u = x_{5n+5}$, $v = x_{5n+6}$, $w = x_{5n+7}$, $\bar{x} = x_{5n+8}$, $\bar{y} = x_{5n+9}$.

Each of these transformations yields a convergence theorem for the continued fraction ξ in accordance with the remarks in §3. For example, K' gives the following theorem.

The continued fraction ξ converges if the following inequalities hold:

$$\left|\frac{x_2}{1-x_3}\right| \leqq \frac{1}{4}; \quad \left|\frac{x_{5n+3}}{x_{5n+3}-1}\right| \leqq \frac{1}{4}, \quad \left|\frac{1}{x_{5n+4}}\right| \leqq \frac{1}{4},$$
$$(n = 0, 1, 2, \cdots),$$

$$\left|\frac{x_{5n}}{x_{5n}-1}\right| \leqq \frac{1}{4}, \quad \left|\frac{x_{5n+1}}{1-x_{5n+2}}\right| \leqq \frac{1}{4}, \quad \left|\frac{x_{5n+2}}{1-x_{5n+3}}\right| \leqq \frac{1}{4},$$
$$(n = 1, 2, 3, \cdots),$$

where $x_1, x_2, x_3, \cdots$ are real or complex numbers.

NORTHWESTERN UNIVERSITY

ON THE MINIMIZING PROPERTY OF THE HARMONIC FUNCTION

BY E. J. MCSHANE

1. *Introduction.* Let D be a bounded connected open set and D^* its boundary and let $\overline{D} = D + D^*$. It is well known † that if the function $u(x, y)$ be harmonic on D, then $u(x, y)$ minimizes the Dirichlet integral

$$I[f] = \int\int_D [f_x^2 + f_y^2]\,dx\,dy$$

in the class of all functions $f(x, y)$ possessing piecewise continuous partial derivatives f_x and f_y and coinciding with $u(x, y)$ on the boundary D^*. But in certain recent discussions of the problem of Plateau‡ essential use is made of a generalization of this theorem; it was necessary to know that the harmonic function $u(x, y)$ minimizes $I[f]$ in a larger class of functions than those with continuous derivatives. It has been suggested that a proof of this fact should be published; the present note carries out the suggestion. The method of proof is similar to that due to Lebesgue. As an application, a theorem is proved which is of some interest in the theory of curved surfaces.

The functions with which we shall be concerned are those which are, as we shall say, *absolutely continuous by sections* (abbreviated a.c.s.). A function $v(x, y)$, defined and continuous on a bounded open set D, will be said to be a.c.s. on D if it satisfies the following conditions:

(1 a) *for almost all values y_0 of y the function $v(x, y_0)$ is absolutely continuous on each interval of the line $y = y_0$ lying in D;*

(1 b) *for almost all values x_0 of x it is absolutely continuous on each interval of the line $x = x_0$ lying in D.*

---

† H. Lebesgue, Société Mathématique de France, Comptes Rendus, (1913), p. 48. Hurwitz-Courant, *Funktionentheorie*, 2d ed., p. 424.

‡ E. J. McShane, *Parametrizations of saddle surfaces*, etc., Transactions of this Society, vol. 35 (1933), pp. 716–733. T. Radó, *The problem of Plateau*, vol. II, No. 3, of the Ergebnisse der Mathematik und ihre Grenzgebiete, p. 99.

With this terminology we state our theorem on the minimizing property of the harmonic function.

THEOREM 1. *If* (a) *u is continuous on $\overline{D}$ and harmonic on D,* (b) *v is continuous on $\overline{D}$ and a.c.s. on D,* (c) *$v(x, y)=u(x, y)$ on D^*, then $I[v]\geqq I[u]$, the sign of equality holding only if $I[u]=\infty$ or $v\equiv u$.*

Before proceeding to the proof of the theorem, we first observe that if $f(x, y)$ be a.c.s. on D, its derivatives v_x and v_y are defined almost everywhere on D and are measurable where defined. Where they are undefined, we assign them the value 0. We shall denote by X the set of all values x_0 of x such that (a) the line $x=x_0$ has points in common with D, (b) on each interval of the line $x=x_0$ lying in D the function $v(x_0, y)$ is absolutely continuous in y. The set Y is defined analogously. For each x_0 of X the line $x=x_0$ has in common with D a finite or denumerable set of intervals; these we denote by $\delta_1(x_0)$, $\delta_2(x_0)$, $\cdots$. The intervals $\delta_i(y_0)$ are analogously defined. As a first step in the proof of Theorem 1 we establish the following lemma.

LEMMA. *If $u(x, y)$ is harmonic on an open set containing $\overline{D}$, and v is continuous on $\overline{D}$ and a.c.s. on D, and $v(x, y)=u(x, y)$ on D^*, then $I[v]\geqq I[u]$.*

If $I[v]=\infty$, the conclusion holds. Otherwise let us define

$$\phi(x, y)=v(x, y)-u(x, y).$$

This function is continuous on $\overline{D}$ and a.c.s. on D, and

$$(1)\qquad I[v]=I[u]+I[\phi]+2\int\int_D (u_x\phi_x+u_y\phi_y)dx\,dy;$$

here all integrals are well defined and finite. We may write

$$\int\int_D u_x\phi_x\,dx\,dy=\int_Y\left\{\int_{\Sigma\delta_i(y)} u_x\phi_x\,dx\right\}dy$$

$$=\int_Y\left\{\Sigma\int_{\delta_i(y)} u_x\phi_x\,dx\right\}dy.$$

In the last integral we integrate by parts, remembering that ϕ vanishes at each end of each $\delta_i(y)$. We thus obtain

$$\iint_D u_x\phi_x\,dx\,dy = -\int_Y \Big\{ \sum \int_{\delta_i(y)} u_{xx}\phi\,dx \Big\} dy$$

$$= -\iint_D u_{xx}\phi\,dx\,dy; \tag{2}$$

the last reduction is allowable because u_{xx} and ϕ are both continuous and bounded on $\overline{D}$. In a like manner,

$$\iint_D u_y\phi_y\,dx\,dy = -\iint_D u_{yy}\phi\,dx\,dy.$$

Adding and remembering that u is harmonic, we find that the last integral of equation (1) vanishes. Since $I[\phi] \geqq 0$, the lemma is established.

Proceeding to the proof of the theorem, we subdivide D into the sets D_0, D', D'', on which the respective relations $v=u$, $v>u$, $v<u$ hold. Also, for every $\epsilon>0$, we define D_ϵ' and D_ϵ'' to be the subsets of D on which $v>u+\epsilon$ and $v<u-\epsilon$, respectively. The Dirichlet integrals over these sets will be distinguished by the corresponding affixes; for example, $I_\epsilon'(u)$ is the integral of $u_x^2+u_y^2$ over D_ϵ'.

We first observe that $I_0[v]=I_0[u]$; in fact, $u_x=v_x$ and $u_y=v_y$ for almost all points of D_0. For let E be the set on which $u-v=0$, the derivatives u_x and v_x are defined, and $u_x-v_x\neq 0$. This set is measurable; hence to prove that its measure is zero, we need only show that for almost all y_0 the part of E lying on the line $y=y_0$ has linear measure zero. But for every y_0 the points of E belonging to the line $y=y_0$ form an isolated set and so are enumerable; for at each such point we have $u-v=0$, while $\partial(u-v)/\partial x$ exists and is not zero. Therefore $m(E)=0$.

As ϵ tends to 0, the sets D_ϵ', D_ϵ'' tend to D', D'', respectively; hence $I_\epsilon'[u]\to I'[u]$ and $I_\epsilon''[u]\to I''[u]$. Thus if h be any number less than $I[u]$, we can choose ϵ small enough so that $I_\epsilon'[u]+I_\epsilon''[u]+I_0[u]>h$. The set D_ϵ' lies, with its boundary $D_\epsilon'^*$, in D; and on $D_\epsilon'^*$ we have $v-\epsilon=u$. Hence by our lemma $I_\epsilon'[v]=I_\epsilon'[v-\epsilon]\geqq I_\epsilon'[u]$, and similarly $I_\epsilon''[v]\geqq I_\epsilon''[u]$. Thus

$$I[v] \geqq I_0[v] + I_\epsilon'[v] + I_\epsilon''[v] \geqq I_0[u] + I_\epsilon'[u] + I_\epsilon''[u] > h.$$

This being true for every $h<I[u]$, it follows that $I[v]\geqq I[u]$, as was to be proved.

It remains to be shown that if $I[u]$ is finite and $I[v]=I[u]$, then $v\equiv u$. In this case we consider the function

$$I[u+\lambda\phi] = I[u] + \lambda^2 I[\phi] + 2\lambda \iint_D (u_x\phi_x + u_y\phi_y)dx\,dy,$$

where $\phi=v-u$. On setting $\lambda=1$, we find that the last integral on the right has the value $-I[\phi]/2$. Setting $\lambda=1/2$, the equation becomes $I[u+\phi/2]=I[u]-I[\phi]/4$. Therefore, if $I[\phi]>0$, the a.c.s. function $u+\phi/2$ has a Dirichlet integral less than that of u, which we have already shown impossible; hence $I[\phi]=0$. This implies that $\phi_x=0$ almost everywhere in D. Thus for almost all y_0 of Y the equation $\phi_x(x, y_0)=0$ holds for almost all x in $\sum\delta_i(y_0)$; integrating and remembering that $\phi=0$ at the ends of each $\delta_i(y_0)$, we find $\phi(x, y)=0$ for almost all (x, y) in D. Since ϕ is continuous, $\phi=u-v$ vanishes identically on D, completing the proof of the theorem.

2. *Discussion of a.c.s. Functions.* As yet we have not shown that the class of a.c.s. functions includes the class of functions with piecewise continuous derivatives and finite Dirichlet integral. Suppose, then, that $v(x, y)$ and its derivatives v_x and v_y are continuous on D, and $I[v]$ is finite; we state that v is a.c.s. For the finiteness of $I[v]$ implies that $|v_x(x, y_0)|$ is summable over $\sum\delta_i(y_0)$ for almost all y_0. Since v_x is continuous in D, v is absolutely continuous over every closed interval contained in $\delta_i(y_0)$; and this, with the summability of $v_x(x, y_0)$, implies that v is absolutely continuous on $\delta_i(y_0)$. A like argument applies to $v(x_0, y)$ for almost all x_0. In particular, if $u(x, y)$ is harmonic on D and $I[u]$ is finite, then $u(x, y)$ is a.c.s.

If v is continuous on D and $I[v]$ is finite, and D can be subdivided into a finite number of subsets, each bounded by a finite number of rectifiable simple arcs such that v_x and v_y are continuous on the interior of each subset, then v is a.c.s. For, if we except a set of x_0 of measure zero, the line $x=x_0$ intersects the boundary curves in a finite number of points, so that each $\delta_i(x_0)$ is subdivided into a finite number of sub-intervals. Again except on a set of measure zero, the function $v(x_0, y)$ is absolutely continuous on each of these subintervals by the preceding paragraph, and being also continuous on $\delta_i(x_0)$ it must be absolutely continuous on $\delta_i(x_0)$. A like argument holds for almost all y_0.

3. *Invariance Properties.* The definition of the property a.c.s. seems highly artificial, since it is stated in terms of a particular coordinate system and has no self-evident invariance properties even under rotations of the axes. But this defect is only apparent; in fact, we shall now apply Theorem 1 to show that if a function is a.c.s. and has a finite Dirichlet integral, it retains these properties under all conformal transformations, the Dirichlet integral being unchanged in value. We shall prove the following theorem.

THEOREM 2. *Let the transformation*

$$x = x(x', y'), \quad y = y(x', y') \tag{3}$$

map the bounded open set D' conformally on the bounded open set D and map $\overline{D}'$ topologically on $\overline{D}$, and let the function $v'(x', y')$ be a.c.s. on D' and have a finite Dirichlet integral

$$I'[v'] \equiv \iint_{D'} [v'^2_{x'} + v'^2_{y'}]dx'dy'.$$

Then its transform $v(x, y) = v'(x'(x, y), y'(x, y))$ is continuous on $\overline{D}$ and a.c.s. on D, and $I[v] = I'[v']$.

By drawing lines parallel to the x- and y-axes we subdivide the (x, y)-plane into squares of side 2^{-n}, where n is an arbitrary positive integer. We obtain the $(n+1)$th subdivision by subdividing each square of the nth. Those squares which are (with their boundaries) interior to D we call $q_1, q_2, \cdots, q_m$; the remainder of D is a set r, bounded in part by D^* and in part by line segments. In the q_j and in r we construct the harmonic functions which coincide with $v(x, y)$ on the boundaries of the subsets. Thus we have defined a function $u_n(x, y)$ which is continuous on D and harmonic on the interiors of the q_j and r. The transformation (3) carries $r, q_1, q_2, \cdots, q_m$ into subsets of D', which we call $r', q_1', q_2', \cdots, q_m'$, respectively; and it carries u_n' into a function $u_n'(x', y')$ which is continuous on D' and harmonic on the interior of each q_j' and r' and which coincides with v' on the boundary of each subset. Hence, by Theorem 1, the Dirichlet integral of v_n' over each subset is at most equal to the Dirichlet integral of u' over that set, so on adding these integrals we have $I'[u_n'] \leqq I'[v']$. But in each set q_i, r, the map-

ping (3) leaves the Dirichlet integral of the harmonic function u unchanged; hence $I[u_n]=I'[u_n']\leqq I'[v']$. Let us consider a point (x_0, y_0) of D. For all sufficiently large n, it belongs to some square, say $q_{i(n)}$, of the nth subdivision. Since the diameters of the squares approach zero, the greatest and least of the values of $v(x, y)$ on the boundary of $q_{i(n)}$ tend to $v(x_0, y_0)$; and since u_n is harmonic on the square, the value $u_n(x_0, y_0)$ lies between these two extreme values of v, so that $\lim u_n(x, y)=v(x,y)$ at each point of D. This equation continues to hold on D^*, where every u_n coincides with v. Finally, by §2, each $u_n(x, y)$ is a.c.s. on D.

We are now in a position to repeat the demonstration on pages 719–720 of my previously cited paper; we need only replace the word "uniformly" by "everywhere" and replace the interated integrals taken first with respect to v between $-(1-u^2)^{1/2}$ and $(1-u^2)^{1/2}$ and then with respect to v over CE, by iterated integrals taken first over $\sum\delta_i(x)$ and then over X. We thus find that $v(x, y)$ is a.c.s., and that

$$I[v] \leqq \liminf I[u_n] = \liminf I'[u_n'] \leqq I'[v'].$$

Interchanging the roles of v and v' shows that $I'[v']\leqq I[v]$; hence the Dirichlet integrals are equal, and the theorem is established. This theorem has an application in the theory of curved surfaces. Among all continuous surfaces

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \qquad ((u, v) \text{ on } B),$$

where B is a Jordan region, a particularly interesting class is that for which

(α) the functions $x(u, v)$, $y(u, v)$, $z(u, v)$ are a.c.s. on B,
(β) $E+G$ is summable, that is, $I[x]$, $I[y]$, $I[z]$ are finite.

In particular, if the region B can be taken to be a circle, α and β remaining satisfied, we say (with Morrey) that the surface is of type L_2. It is clearly more convenient to have a circle to deal with than to have a general Jordan region. But we can now show that every surface having a representation satisfying (α), (β) is of type L_2. We need only map B conformally on the unit circle; the transformed functions $\bar{x}(u, v)$, $\bar{y}(u, v)$, $\bar{z}(u, v)$, $u^2+v^2\leqq 1$, then satisfy (α) and (β) because of Theorem 2.

PRINCETON UNIVERSITY

SIMPLIFICATION OF THE WHITEHEAD-HUNTINGTON SET OF POSTULATES FOR THE ALGEBRA OF LOGIC

BY A. H. DIAMOND*

1. *Introduction.* In 1916, B. A. Bernstein† obtained a reduction in the number of postulates of the Whitehead-Huntington set‡ for the algebra of logic. In this paper I secure a further reduction in the number of the postulates by combining the commutative and distributive laws and then proceed to investigate the postulates which result when the two reductions are effected at once.

I prove the sufficiency of the new postulates for the algebra of logic by showing that the original set can be derived from them. Finally I establish the independence of the new postulates by furnishing proof-systems of a simple arithmetic type.

2. *The New Postulates.* The new postulates have as undefined ideas the undefined ideas of the original set, namely, a *class K* and two *binary operations* $+$ and $\times$. The postulates are the Postulates 1–6 listed below. In Postulates 3 and 4 is implied the condition *if the elements involved and their indicated combinations belong to K.*

POSTULATE 1. *$a+b$ is in K whenever a and b are in K.*

POSTULATE 2. *ab is in K whenever a and b are in K.*

POSTULATE 3. $bc+a=(c+a)(b+a)$.

POSTULATE 4. $(b+c)a=ca+ba$.

POSTULATE 5. *For any element b in K there exists an element b' such that, whatever a is, $a+bb'=a$ and $a(b+b')=a$.*

POSTULATE 6. *There are at least two elements, a and b, in K such that $a\neq b$.*

3. *Sufficiency of the Postulates. Derivation of the Whitehead-Huntington Postulates.* The Whitehead-Huntington postulates leave undefined a *class K* and two *binary operations* $+$ and $\times$,

* American Field Service Fellow.

† See this Bulletin, vol. 22 (1916), pp. 458–459.

‡ See the Transactions of this Society, vol. 5 (1904), pp. 288–309. The set referred to is the first of the three sets which occur in Huntington's paper.

and are the Postulates Ia, Ib, · · · , VI below. In Postulates IIIa–IVb is implied the condition *if the elements involved and their indicated combinations belong to K*, and in V is implied the condition *if the elements z and u of* IIa *and* IIb *exist and are unique.*

POSTULATE Ia. *$a+b$ is in K whenever a and b are in K.*

POSTULATE Ib. *ab is in K whenever a and b are in K.*

POSTULATE IIa. *There is an element z such that $a+z=a$ for every element a.*

POSTULATE IIb. *There is an element u such that $au=a$ for every element a.*

POSTULATE IIIa. $a+b=b+a$.

POSTULATE IIIb. $ab=ba$.

POSTULATE IVa. $a+bc=(a+b)(a+c)$.

POSTULATE IVb. $a(b+c)=ab+ac$.

POSTULATE V. *For every element a there is an element a' such that $a+a'=u$ and $aa'=z$.*

POSTULATE VI. *There are at least two elements, a and b, in K such that $a\neq b$.*

It is seen that Postulates 1, 2, 6 are precisely the Postulates Ia, Ib, VI, respectively, of the original set. Further, Bernstein has shown* that IIa, IIb, V of the original set follow from Postulates 1, 2, and 5. Consequently, in order to prove the sufficiency of Postulates 1–6 for the algebra of logic, it remains to derive from the postulates 1–6 Postulates IIIa–IVb of the original set. In the proofs I give below I assume, in accordance with Bernstein's results, that IIa and IIb have already been proved.

PROOF OF IIIa. In 4 put $a=u$, using 1, 2, and IIb.

PROOF OF IIIb. In 3 put $a=z$, using 1, 2, and IIa.

PROOF OF IVa. By 1, 2, 3, IIIa, and IIIb.

PROOF OF IVb. By 1, 2, 4, IIIa, and IIIb.

It can easily be verified that Postulates 3 and 4 can be derived from IIIa–IVb. It thus becomes apparent, when one also takes into account the results obtained by Bernstein, that Postulates 1–6 are equivalent to Postulates Ia–VI of the original set.

4. *Independence of the Postulates.* I establish the independence of Postulates 1–6 by giving the following examples of proof-

* See this Bulletin, loc. cit.

systems, each of which violates the like-numbered postulate and satisfies all the other postulates of the set. These examples are all *arithmetic* systems, the elements being the numbers 0 and 1. The symbol $f(a, b)$ (mod 2) in an arithmetic system denotes the *least positive residue modulo* 2 obtained from $f(a, b)$ by rejecting multiples of 2. The operations $+$ and $\times$ are to be interpreted as the operations of ordinary arithmetic when they occur in the modular expression, otherwise they are to be interpreted as logical addition and logical multiplication.

EXAMPLE	K	$a+b$	ab
1	0, 1	$a+b+0/ab+1$ (mod 2)	ab (mod 2)
2	0, 1	ab (mod 2)	$a+b+0/ab+1$ (mod 2)
3	0, 1	$a+b$ (mod 2)	ab (mod 2)
4	0, 1	ab (mod 2)	$a+b$ (mod 2)
5	0, 1	ab (mod 2)	ab (mod 2)
6	Null		

PARIS, FRANCE

ON ANALYTIC FUNCTIONS WITH POSITIVE IMAGINARY PARTS

BY J. L. DOOB* AND B. O. KOOPMAN

The purpose of this paper is to give an integral representation of a function analytic in a half-plane, and with positive imaginary part there. This can be used to obtain in a simple way the well known analytic representation of the resolvent of a self-adjoint transformation in abstract Hilbert space.

THEOREM. *Let $\phi(l)$ be a function analytic for $\Im(l)>0$.† If*

$$\Im[\phi(l)] \geqq 0, \qquad \limsup_{t\to\infty} |t\Im[\phi(it)]| < \infty, \tag{1}$$

for t real, $t>0$, then there is a uniquely determined monotone non-decreasing function $\alpha(\lambda)$, defined for $-\infty<\lambda<\infty$, satisfying

* National Research Fellow.

† If ξ is any complex number, $\Re(\xi)$ and $\Im(\xi)$ will be used to denote its real and imaginary parts, and $\bar{\xi}$ its conjugate.

$$\lim_{\lambda \to -\infty} \alpha(\lambda) = 0, \qquad \alpha(\lambda + 0) = \alpha(\lambda), \tag{2}$$

and

$$\alpha(\lambda) \leqq \limsup_{t \to \infty} \left| t\Im[\phi(it)] \right|, \quad (-\infty < \lambda < \infty), \tag{3}$$

and such that

$$\phi(l) = \int_{-\infty}^{\infty} \frac{d\alpha(\lambda)}{\lambda - l} + c, \tag{4}$$

*where c is a real constant.**

The function

$$\phi\left(-i\frac{z+1}{z-1}\right)$$

is analytic for $|z|<1$ and $\Im(\phi) \geqq 0$ there. Then by a theorem of Herglotz,† there is a non-decreasing function $\alpha_1(\theta)$ defined for $0 \leqq \theta \leqq 2\pi$, such that, if $z = re^{i\psi}$,

$$\Im\phi\left(-i\frac{z+1}{z-1}\right) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1-2r\cos(\theta-\psi)+r^2}\, d\alpha_1(\theta). \tag{5}$$

If $\alpha_2(\theta) = \alpha_1(\theta+0) - \alpha_1(0+0)$ for $0 \leqq \theta < 2\pi$, $\alpha_2(2\pi) = \alpha_1(2\pi-0) - \alpha_1(0+0)$, this becomes

$$\begin{aligned} \Im\phi\left(-i\frac{z+1}{z-1}\right) &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1-2r\cos(\theta-\psi)+r^2}\, d\alpha_2(\theta) \\ &\quad + \frac{c_1}{2\pi}\,\frac{1-r^2}{1-2r\cos\psi+r^2}, \end{aligned} \tag{6}$$

where $c_1 = \alpha_1(2\pi) - \alpha_1(2\pi-0) + \alpha_1(0+0) - \alpha_1(0)$, and is evidently

* See results of R. Nevanlinna, Annales Academiae Scientiarum Fennicae, Series A, vol. 18 (1922), No. 5; W. Cauer, this Bulletin, vol. 38 (1932), p. 715, especially the footnote; Marshall Stone, *Linear Transformations in Hilbert Space*, Theorem 10.36; J. Wolff and F. de Kok, Bulletin de la Société Mathématique de France, vol. 50 (1932), pp. 221–227; S. Warschawski, Mathematische Annalen, vol. 107 (1932), pp. 1–15, who discusses integrals in the circle with the origin as center and of radius 1 which correspond to our integrals in the upper half-plane.

† Leipziger Berichte, Mathematische-Physikalische Klasse, vol. 63 (1911), pp. 508–511.

not negative. Moreover the function $\alpha_2(\theta)$ is uniquely determined by $\Im(\phi)$; and $\alpha_1(\theta)$ is determined on a set of measure 2π, to within a constant,* which fixes $\alpha_2(\theta)$ uniquely. If we change θ, z, to λ and l, respectively, by

$$\lambda = -i\frac{e^{i\theta}+1}{e^{i\theta}-1} = -\operatorname{ctn}\frac{\theta}{2}, \qquad l = -i\frac{z+1}{z-1},$$

(6) becomes, if $\alpha_2(\theta)=\alpha_3(\lambda)$,

$$(7)\quad \Im[\phi(l)] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\Im(l)}{|\lambda-l|^2}(1+\lambda^2)d\alpha_3(\lambda) + \frac{c_1}{2\pi}\Im(l).$$

Now if $l=it$, $t>0$, this becomes

$$(8)\qquad \Im[\phi(it)] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{t}{\lambda^2+t^2}(1+\lambda^2)d\alpha_3(\lambda) + \frac{c_1 t}{2\pi},$$

and it is seen at once that (1) implies that $c_1=0$ and that the integral $(1/\pi)\int_{-\infty}^{\infty}(1+\lambda^2)d\alpha_3(\lambda)$ is finite and not greater than† $\limsup_{t\to\infty}|t\Im[\phi(it)]|$. If, then, $\alpha(\lambda)$ is defined by

$$(9)\qquad \alpha(\lambda) = \frac{1}{\pi}\int_{-\infty}^{\lambda}(1+\lambda^2)d\alpha_3(\lambda),$$

it is seen to satisfy (2) and (3). Substituting in (7), we find

$$(10)\qquad \Im[\phi(l)] = \int_{-\infty}^{\infty}\frac{\Im(l)}{|\lambda-l|^2}d\alpha(\lambda),$$

---

* For (5) is equivalent to

$$\phi_1(z) = \frac{1}{i}\phi\left(-i\frac{z+1}{z-1}\right) = \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}d\alpha_1(\theta) + id,$$

where d is a real constant, from which it is seen at once that

$$\phi_1^{(n)}(0) = \frac{n!}{2\pi}\int_0^{\pi}e^{-ni\theta}d\alpha_1(\theta),\ n\geqq 1,\ \Im\phi(i) = \frac{1}{2\pi}\int_0^{2\pi}d\alpha_1(\theta).$$

This determines $\alpha_1(\theta)$ to within a constant on a set of measure 2π, since it is a real function.

† This is the crucial step of the proof and since the same conclusions can be drawn if there is a sequence $\{\rho_n e^{i\tau_n}\}$, $(0\leq\tau_n<\pi)$, such that $\liminf_{n\to\infty}\tau_n(\pi-\tau_n)>0$, and such that $\limsup_{n\to\infty}|\rho_n e^{i\tau_n}\Im[\phi(\rho_n e^{i\tau_n})]|<\infty$, the lemma can be generalized accordingly.

which is equivalent to

$$\phi(l) = \int_{-\infty}^{\infty} \frac{d\alpha(\lambda)}{\lambda - l} + c, \tag{11}$$

where c is a real constant. The function $\alpha(\lambda)$, monotone non-decreasing and satisfying (2), is thereby uniquely determined. For otherwise, retracing steps, the fact that $\alpha_2(\theta)$ is uniquely determined would be contradicted. The representation (10) shows that unless $\phi(l) \equiv c$, $t\Im\phi(it)$ increases monotonely with t and that $\lim_{\lambda \to -\infty} \alpha(\lambda) = \lim_{t \to -\infty} t\Im[\phi(it)]$.

COROLLARY. *If $\phi(l)$ is a function analytic for $\Im(l) \neq 0$ which can be put in the form* (4) *with $c = 0$, where $\alpha(\lambda)$ is a complex-valued function satisfying* (2), *whose real and imaginary parts are of bounded variation over the interval $-\infty < \lambda < \infty$, then $\alpha(\lambda)$ is thereby uniquely determined.*

This has been shown already (without supposing that $\phi(l)$ is defined for $\Im(l) < 0$) when $\alpha(\lambda)$ is supposed real and non-decreasing. The case when $\alpha(\lambda)$ is only supposed real can be reduced to this case, since any real function of bounded variation can be expressed as the difference between two monotone non-decreasing functions. In the general case we only have to show that $\phi(l) \equiv 0$ implies that $\alpha(\lambda) \equiv 0$. Suppose, then, that $\phi(l) \equiv 0$, and define $\phi_1(l)$, $\phi_2(l)$ by

$$\phi_1(l) = \int_{-\infty}^{\infty} \frac{d\Re\alpha(\lambda)}{\lambda - l}, \qquad \phi_2(l) = \int_{-\infty}^{\infty} \frac{d\Im\alpha(\lambda)}{\lambda - l}.$$

Then

$$\phi_1(l) + i\phi_2(l) \equiv 0, \qquad \Im(l) \neq 0. \tag{12}$$

Substituting $\bar{l}$ for l, $\phi_1(\bar{l}) + i\phi_2(\bar{l}) \equiv 0$, $\Im(l) \neq 0$. But it is evident that $\phi_j(\bar{l}) = \overline{\phi_j(l)}$, $(j = 1, 2)$, so that $\overline{\phi_1(l) + i\phi_2(l)} \equiv 0$, or, if we take the conjugate,

$$\phi_1(l) - i\phi_2(l) \equiv 0, \qquad \Im(l) \neq 0. \tag{13}$$

Equations (12) and (13) show that $\phi_1(l) \equiv \phi_2(l) \equiv 0$. We have already seen that this implies that $\Re\alpha(\lambda) \equiv \Im\alpha(\lambda) \equiv 0$, as was to be proved.

These results can be applied to the theory of self-adjoint

transformations in Hilbert space as follows. Let T be a self-adjoint transformation* in Hilbert space $\mathfrak{H}$, and let R_l be the inverse of the transformation $T-lI$, where $\Im(l)\neq 0$, and I is the identical transformation. Then if f is an arbitrary element of $\mathfrak{H}$, $[R_l f, f]$ is analytic for $\Im(l)\neq 0$ and $\Im[R_l f, f]$ has the same sign† as $\Im(l)$. By the theorem proved in this paper there is then an integral representation of $[R_l f, f]$:

$$[R_l f, f] = \int_{-\infty}^{\infty} \frac{d\alpha(\lambda; f)}{\lambda - l}. \tag{14}$$

If f and g are arbitrary elements of $\mathfrak{H}$,

$$\begin{aligned}[R_l f, g] = {} & \frac{1}{4}[R_l(f+g), f+g] - \frac{1}{4}[R_l(f-g), f-g] \\ & + \frac{i}{4}[R_l(f+ig), f+ig] - \frac{i}{4}[R_l(f-ig), f-ig],\end{aligned} \tag{15}$$

which leads to an integral representation of $[R_l f, g]$,

$$[R_l f, g] = \int_{-\infty}^{\infty} \frac{d\beta(\lambda; f, g)}{\lambda - l}, \tag{16}$$

where $\beta(\lambda; f, g)$ is a complex-valued function whose real and imaginary parts are of bounded variation over the infinite interval. It can then be shown that $\beta(\lambda; f, g) = (E(\lambda)f, g)$, where the transformations $E(\lambda)$ form a resolution of the identity.‡

COLUMBIA UNIVERSITY

---

* The notation used will be that in Stone, *Linear Transformations in Hilbert Space*, Colloquium Publications of this Society, vol. 15.

† Stone, loc. cit., Chapters 4 and 5.

‡ Carrying out the details of this would mean following the treatment given by Stone except that the considerations of Chapter 5, §2, would be omitted, Theorem 5.4 being obtained directly from the theorem of this paper, as just explained. This treatment is similar in spirit to those given by Bochner, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1933, pp. 371–376, and by A. Khintchine, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 567–573, who also separate the purely analytic facts out of certain results in Hilbert space.

METRICALLY TRANSITIVE POINT TRANSFORMATIONS*

BY M. H. MARTIN

1. *Introduction.* Let

$$(1) \qquad T^{(j)}: \quad \xi_i^{(j)} = \xi_i^{(j)}(x_1, x_2, \cdots, x_n), \qquad (i = 1, 2, \cdots, n;\ j = 1, 2, \cdots),$$

denote a denumerable set of point transformations in n-dimensional euclidean space, the transformations of which leave a certain p-dimensional $(p \leqq n)$ region R (region $\equiv$ open, connected point set) of this space invariant. The case where R is the entire n-dimensional space is not excluded.

The set of transformations (1) *will be said to be metrically transitive with respect to* R, *if the complement set* (*with respect to* R) *of every non-zero subset of* R *that is invariant under each transformation of the set* (1) *is a zero set.*† *If the transformations of the set* (1) *form a group, this group is said to be metrically transitive with respect to* R.

If the set (1) consists of a single transformation T and if $p = 2$, the above definition is sensibly the same as that given by Birkhoff and Smith‡ for metrically transitive surface transformations. If the transformations of the set (1) are composed of the iterations of a single transformation T and its inverse T^{-1}, together with the identity transformation, the transformations of the set form an infinite cyclic group which is metrically transitive with respect to a certain region if, and only if, one of the elements of this group, exclusive of the identity transformation, is metrically transitive with respect to this region.

---

* This paper was written while the author was a National Research Fellow at Harvard University, and was subsequently revised.

† *Non-zero set* $\equiv$ a set of positive p-dimensional Lebesgue measure. *Zero set* $\equiv$ a set of zero p-dimensional Lebesgue measure. In this paper the word measure refers to Lebesgue measure; the measure of a measurable set S is denoted by mS.

‡ G. D. Birkhoff and P. A. Smith, *Structure analysis of surface transformations*, Journal de Mathématiques, vol. 7 (1928), p. 365.

The principal result of this paper is that the elliptic modular group

$$\zeta = \frac{\alpha z + \beta}{\gamma z + \delta}, \tag{2}$$

(where α, β, γ, and δ are integers for which $\alpha\delta - \beta\gamma = 1$ and ζ, z complex variables) is metrically transitive with respect to the real axis. Here $n=2$, $p=1$, and R is the real axis. From the metrical transitivity of this group there follow several results in regard to: (a) the covering of any given point set on the real axis by a denumerable set of point sets,* (b) the distribution of equivalent points [that is, points which may be transformed into one another by transformations of the group (2)] on the real axis, (c) the measure of the sets of Liouville and Hurwitz† numbers on the real axis. In conclusion we consider a very simple metrically transitive set of point transformations and point out an application of our results to a metrically transitive system previously considered by Birkhoff and Smith,‡ and others.

2. *Metrical Transitivity With Respect to the Real Axis.* In order to demonstrate that the group (2) is metrically transitive with respect to the real axis, it is convenient to consider directly the group of transformations

$$\xi = \frac{\alpha x + \beta}{\gamma x + \delta}, \tag{3}$$

(where α, β, γ, and δ are integers for which $\alpha\delta - \beta\gamma = 1$ and ξ, x are real variables) of the real axis into itself, induced on it by the group (2). Let S be any non-zero set on the x-axis invariant under all the transformations of the group (3). Since this group contains the translations $\xi = x \pm n$, $(n = 0, 1, \cdots)$, the comple-

* For other results in this connection see H. Rademacher, *Eineindeutige Abbildungen und Messbarkeit*, Monatshefte für Mathematik und Physik, vol. 27 (1916), pp. 184–195.

† A number is said to be a Hurwitz number if its development into a regular continued fraction is a Hurwitz continued fraction. See O. Perron, *Die Lehre von den Kettenbrüchen*, 1913, pp. 126–127.

‡ Loc. cit., p. 368. See also G. D. Birkhoff, *Probability and physical systems*, this Bulletin, vol. 38 (1932), p. 371.

mentary set of S on the x-axis may be shown to be a zero set by demonstrating that the part S_1 of S contained in the unit interval $0<x<1$ necessarily has measure 1; and it is to the proof of this fact that we now turn.

3. *Measure of S_1 on the Unit Interval.* Let ξ_1 be *any* irrational number in the unit interval $0<\xi<1$ and denote by $[0, p_1, p_2, \cdots, p_j, \cdots]$ its development into an unending regular continued fraction.* We consider the subintervals $I^{(n)}$ of the unit interval $0<\xi<1$ defined by

$$(4)\quad I^{(n)}:\quad [0, p_1, p_2, \cdots, p_{2n}] < \xi < [0, p_1, p_2, \cdots, p_{2n}, 1], \qquad (n = 1, 2, \cdots),$$

each of which contains the point ξ_1. From the theory of continued fractions it is known that these intervals are the transforms of the unit interval $0<x<1$ by the transformations $\xi = [0, p_1, p_2, \cdots, p_{2n}, 1:x]$, or

$$(5)\quad T^{(n)}: \xi = \frac{P_{2n}:x + P_{2n-1}}{Q_{2n}:x + Q_{2n-1}} = \frac{P_{2n-1}x + P_{2n}}{Q_{2n-1}x + Q_{2n}}, \quad (n = 1, 2, \cdots),$$

and hence are identical with the intervals

$$(4)'\quad I^{(n)}:\qquad \frac{P_{2n}}{Q_{2n}} < \xi < \frac{P_{2n-1} + P_{2n}}{Q_{2n-1} + Q_{2n}}, \qquad (n = 1, 2, \cdots).$$

Here P_{2n-1}, P_{2n}, Q_{2n-1}, and Q_{2n} are positive integers calculated from the recursion formulas

$$P_0 = 0, P_1 = 1, \cdots, P_j = p_jP_{j-1} + P_{j-2},$$
$$Q_0 = 1, Q_1 = p_1, \cdots, Q_j = p_jQ_{j-1} + Q_{j-2}, \quad (j = 2, 3, \cdots, 2n).$$

They satisfy the following inequalities:

$$(6)\quad \text{(a) } P_{2n-1}Q_{2n} - P_{2n}Q_{2n-1} = 1;\ \text{(b) } Q_{2n} \geqq 2n;\ \text{(c) } Q_{2n-2} < Q_{2n}; \qquad (n = 1, 2, \cdots).$$

Hence the lengths $mI^{(n)}$ of the intervals $I^{(n)}$ are given by

$$(7)\quad mI^{(n)} = \frac{P_{2n-1} + P_{2n}}{Q_{2n-1} + Q_{2n}} - \frac{P_{2n}}{Q_{2n}} = \frac{1}{Q_{2n}(Q_{2n-1} + Q_{2n})} < \frac{1}{2n}.$$

* For the notation and properties of continued fractions used in this paper, see O. Perron, op. cit.

From (6a) it follows that the transformations $T^{(n)}$ in (5) belong to the group (3). Since the transformation $T^{(n)}$ transforms the unit interval $0<x<1$ into the interval $I^{(n)}$, and, since ξ_1 is contained in each of the intervals $I^{(n)}$, it follows from (7) that the unit interval $0<x<1$ is transformed by the transformations $T^{(n)}$ into a sequence $\{I^{(n)}\}$ of intervals closing down on ξ_1.

We observe that the transformation $T^{(n)}$ can be written in the form

$$(8)\quad T^{(n)}: \qquad \xi = \frac{P_{2n}}{Q_{2n}} + \int_0^x \phi_n(x)dx, \qquad (n = 1, 2, \cdots),$$

where

$$(9)\quad \phi_n(x) = (Q_{2n-1}x + Q_{2n})^{-2} > 0, \quad (0 \leqq x \leqq 1; n = 1, 2, \cdots).$$

From (6a) we conclude that $\phi_n(x)$ is continuous in this interval and hence summable in any subinterval of it. The transformation $T^{(n)}$ is therefore an absolutely continuous function of x in the interval $0<x<1$ and any zero set in the interval $0<x<1$ is accordingly transformed by it into a zero set* in the interval $I^{(n)}$.

The set S_1 (see §2) is contained in a set $\overline{S}_1$ which is the inner limiting set of a monotone decreasing sequence of sets of non-overlapping intervals, such that $m\overline{S}_1 = mS_1$. If we denote by $\Sigma_1^{(n)}$ and $\overline{\Sigma}_1^{(n)}$ the sets in the interval $I^{(n)}$ into which $T^{(n)}$ transforms S_1 and $\overline{S}_1$, respectively, we have†

$$(10)\quad m\Sigma_1^{(n)} = m\overline{\Sigma}_1^{(n)} = \int_{\overline{S}_1} \phi_n(x)dx = \int_0^1 f(x)\phi_n(x)dx,$$

where $f(x)$ is the characteristic function of the set $\overline{S}_1$. Since $\phi_n(x)$ is continuous and $f(x)$ a summable function that is $\geqq 0$ in the interval $0 \leqq x \leqq 1$, we may apply the first mean value theorem to obtain

$$(11)\quad \int_0^1 f(x)\phi_n(x)dx = \phi_n(x_1)\int_0^1 f(x)dx = \phi_n(x_1)mS_1,$$

* See, for example, C. Carathéodory, *Vorlesungen über Reelle Funktionen*, 1927, p. 583.

† See, for example, E. W. Hobson, *The Theory of Functions of a Real Variable*, vol. 1, 1921, p. 593.

$(0 \leqq x_1 \leqq 1)$. Now the set S is invariant with respect to the group (3) and hence invariant with respect to the transformations $T^{(n)}$. Accordingly $\Sigma_1^{(n)} \subset S_1$ and the density of S_1 in the interval $I^{(n)}$ is therefore not less than the density of $\Sigma_1^{(n)}$ in this interval. From (7), (9), (10) and (11) we have then

$$(12)\quad \frac{mI^{(n)} \cdot S_1}{mI^{(n)}} \geqq \frac{m\Sigma_1^{(n)}}{mI^{(n)}} = \frac{Q_{2n}(Q_{2n-1} + Q_{2n})}{(Q_{2n-1}x_1 + Q_{2n})^2} mS_1 = \frac{1+\theta}{(1+\theta x_1)^2} mS_1,$$

where $\theta = Q_{2n-1}/Q_{2n}$. From (6c) we obtain $0 < \theta < 1$, and therefore from (12), and our hypothesis $mS_1 > 0$, there results

$$(13)\quad \varliminf_{n\to\infty} \frac{mI^{(n)} \cdot S_1}{mI^{(n)}} \geqq \varliminf_{n\to\infty} \frac{m\Sigma_1^{(n)}}{mI^{(n)}} > 0.$$

From (13) it follows that, if the metrical density of S_1 exists at the point ξ_1, it is positive. Since ξ_1 is any irrational point of the interval $0 < \xi < 1$, the metrical density of S_1 is either positive or non-existent on the set of irrational points in this interval. However, from the theorem of Lebesgue† on the metrical density of measurable sets, we know that the metrical density of S_1 exists, and is equal to unity at all points of S_1, with the possible exception of a component which is a zero set, and is zero at all points of the complement of S_1 with respect to the interval $0 < \xi < 1$, with the possible exception of the points of a component which is a zero set. Therefore $mS_1 = 1$, which was to be proved.

4. *Conclusions.* From the metrical transitivity of the group (2) with respect to the real axis we are able to draw several obvious conclusions. The first result is that every point set on the real axis may be covered, with the possible exception of the points of a component which is a zero set, by a denumerable set of linear sets obtained from any previously given point set of positive measure on the real axis by means of the transformations of the group (2). In order to demonstrate this let S_2 be any set of positive measure on the real axis and denote by S_i, $(i = 1, 2, \cdots)$, the sets obtained from S_2 by the transformations of the group (2). The set $M = \sum_{i=1}^{\infty} S_i$ is obviously invariant under the transformations of the group (2). Since it is of positive measure,

† See, for example, E. W. Hobson. op. cit., p. 181.

and the group (2) is metrically transitive with respect to the real axis, its complementary set is a zero set. Consequently M must cover any set up to a component of measure zero. This last result is very closely related to a result of Rademacher* obtained by use of the covering theorem of Vitali.

Another result is that any set of positive measure on the real axis necessarily contains equivalent points. For, if a set of positive measure exists which contains no equivalent points, it may be divided into two sets, each of positive measure, such that no point of one is equivalent to any point of the other. A set M, constructed from one of these sets as was the set M above, would, however, cover the other up to the points of a zero set, thus leading to a contradiction.

Finally, we observe that, since the sets of Liouville and Hurwitz numbers on the real axis are invariant† under the transformations of the group (2), either they, or their complementary sets with respect to the real axis, are zero sets.

5. *An Example.* A simple example of a set of point transformations which is metrically transitive with respect to a region R is obtained when the transformations of the set (1) are the translations

$$T^{(j)}: \qquad \xi = x + a_j, \qquad (j = 1, 2, \cdots), \tag{14}$$

for which

$$\lim_{n\to\infty} a_j = 0. \tag{15}$$

Here $n=p=1$ and R is the x-axis (that is, the entire 1-dimensional euclidean space). In order to demonstrate the metrical transitivity of (14) with respect to the x-axis, let S be any measurable set invariant under all the transformations of the set (14). Because of (15), the set S is certainly invariant under the transformations of a subset of (14), namely,

$$\xi = x + \alpha_j, \qquad (j = 1, 2, \cdots),$$

in which $\lim_{j\to\infty}\alpha_j=0$ and $\alpha_i>\alpha_j$ if $i<j$. Starting from any point of the x-axis we lay off upon it a denumerable set of intervals,

* J. Rademacher, loc. cit.

† O. Perron, op. cit., p. 127 and p. 141.

each of length α_1 and possessing only their end points in common with one another, which cover the entire x-axis without lacunae. The density of the set S is obviously the same in each of these intervals and consequently the same in any finite segment of the x-axis whose end points coincide with end points of these intervals. This process is then repeated for intervals of length α_2, of length α_3, etc. One obtains in this manner successive subdivisions of the x-axis into intervals, the lengths of which are arbitrarily small and in each of which the set S possesses, *for any given subdivision*, the same density. Let I_1 and I_2 be any two intervals of the x-axis and let $I_1(\alpha_k)$ and $I_2(\alpha_k)$ denote the two subintervals of I_1 and I_2, respectively, which are built up from all those intervals of the kth subdivision that are *entirely* contained in I_1 and I_2. Obviously $\lim_{k\to\infty} mI_1(\alpha_k) = mI_1$ and $\lim_{k\to\infty} mI_2(\alpha_k) = mI_2$. Since the density of the set S is the same in the two intervals $I_1(\alpha_k)$ and $I_2(\alpha_k)$ for any value of k, it is the same in the two intervals I_1 and I_2. Since I_1 and I_2 were arbitrary intervals, the set S is of homogeneous density and either it or its complementary set is of measure zero.*

In the metrically transitive system considered by Birkhoff and Smith it is shown† by a Fourier analysis that if θ and $\bar{\theta}$ are angular variables of period 1, the transformation

$$\bar{\theta} = \theta + \alpha, \qquad (\alpha \text{ irrational}), \tag{16}$$

of the periphery of a circle into itself, is metrically transitive. The same result follows by means of the methods of this section, since (16) is metrically transitive, if the set of transformations

$$\xi = x + n\alpha - [n\alpha], \qquad (n = 1, 2, \cdots),$$

(where $[n\alpha]$ is the greatest integer, in absolute value, for which $|[n\alpha]| < |n\alpha|$) of the x-axis into itself, is metrically transitive with respect to the x-axis. That this is actually so is seen by putting $a_j = j\alpha - [j\alpha]$ in (14) and observing that (15) is fulfilled.

TRINITY COLLEGE

---

* E. Jacobsthal and K. Knopp, *Bemerkungen über die Struktur linearer Punktmengen*, Sitzungsberichte der Berliner Mathematischen Gesellschaft, No. 129 (1915), p. 122.

† G. D. Birkhoff, *Probability and physical systems*, loc. cit.

ON THE LATTICE THEORY OF IDEALS†

BY GARRETT BIRKHOFF

1. *Outline.* The ideals of any ring define, relative to g.c.f. and l.c.m., a combinatorial system having properties which we shall presently define as characterizing *B-lattices.*

In this article we shall first develop some new properties of *B*-lattices as abstract systems; the main results of this part of the work find expression in Theorems 1–5. Then we shall apply this theory and some older results to the ideals of commutative rings R which possess a principal unit l and satisfy the Basis Theorem. In addition to developing the known theory of *einartig* ideals by combinatory methods, we give a necessary and sufficient condition that the *B-lattice* defined by the ideals of R should be isomorphic with a *ring* of point sets in the sense of Hausdorff.‡

2. *Notation*; *Lattice Algebras.* We shall in general use capital letters to denote systems, and small letters for elements. $a \in A$ will mean "a is an element of the system A"; $B \subset A$ will mean "$b \in B$ implies $b \in A$"; $B < A$ will mean $B \subset A$ but $B \neq A$.

By a *lattice algebra* will be meant any system L which satisfies the following postulates:

(L1). Any $a \in L$ and $b \in L$ determine a unique "*join*" $a \cap b \in L$ and a unique "*meet*" $(a, b) \in L$.

(L2). $a \cap b = b \cap a$ and $(a, b) = (b, a)$ for any $a \in L$ and $b \in L$.

(L3). $a \cap (b \cap c) = (a \cap b) \cap c$ and $(a, (b, c)) = ((a, b), c)$ for any $a \in L$, $b \in L$, and $c \in L$.

(L4). $a \cap (a, b) = a$ and $(a, a \cap b) = a$ for any $a \in L$ and $b \in L$.

From (L1)–(L4) follow $a \cap a = (a, a) = a$. Moreover $a \cap b = b$ is equivalent to $(a, b) = a$; in this case we write $a \subset b$ or $b \supset a$, and $a \subset b$ taken with $b \subset c$ implies $a \subset c$. Moreover, $a < b$ means $a \subset b$ but $a \neq b$, while "b covers a" means $a < b$, but that no $x \in L$ satisfies $a < x < b$.

The reader may find it helpful to regard lattices as distorted

† Presented to the Society, March 30, 1934.

‡ Hausdorff, *Mengenlehre*, 1927, p. 77.

Boolean algebras in which $a \cap b$ is substituted for $a+b$, and (a, b) for $a \cdot b$.

The following additional conditions are optional:

(L5). If $a \subset c$, then $a \cap (b, c) = (a \cap b, c)$.

(L6). $(a, b \cap c) = (a, b) \cap (a, c)$ for any $a \in L$, $b \in L$, and $c \in L$.

If a lattice satisfies (L5), it is called a B-lattice; if it satisfies (L6), it is called a C-lattice. Any C-lattice is a B-lattice, and also satisfies $a \cap (b, c) = (a \cap b, a \cap c)$.

3. *Subdirect Decomposition.* We shall consider in §§3–4 only lattices L which have a "*largest*" element j satisfying $a \cap j = j$ for every $a \in L$; such is always the case in applications.†

We shall say that $a \in L$ and $b \in L$ are *coprime* if and only if $a \cap b = j$. We shall say that two sublattices‡ $A \subset L$ and $B \subset L$ are *coprime* if and only if $a \in A$ and $b \in B$ imply $a \cap b = j$. We shall say that the sublattices of a finite or transfinite§ sequence of sublattices $A_1 \subset L, \cdots, A_n \subset L$ are *strongly coprime* if and only if every A_i is coprime with the sublattice generated by‖ the other sublattices of the sequence.

Let $B_1, \cdots, B_n$ be any (finite or transfinite) sequence of lattices, whose largest elements are $j_1, \cdots, j_n$. By an *f-type* vector $[b_1, \cdots, b_n]$, $(b_i \in B_i)$, we mean one in which $b_i = j_i$ except for a *finite* set of subscripts i. By the *subdirect product* $B_1 \hat{x} \cdots \hat{x} B_n = B^*$ of the B_i is meant the lattice whose elements are the f-type vectors just defined, and such that by definition

$$[b_1, \cdots, b_n] \cap [b_1', \cdots, b_n'] = [b_1 \cap b_1', \cdots, b_n \cap b_n'],$$

$$([b_1, \cdots, b_n], [b_1', \cdots, b_n']) = [(b_1, b_1'), \cdots, (b_n, b_n')].$$

B^* is evidently a lattice with largest element $[j_1, \cdots, j_n]$. Further, if B_i^* denotes the sublattice of elements of the form $[j_1, \cdots, j_{i-1}, b_i, j_{i+1}, \cdots, j_n]$ of B^*, then B_i^* is isomorphic with B_i, the lattices $B_1^*, \cdots, B_n^*$ are strongly coprime, and any element of B^* can be expressed as the meet of a finite num-

† In fact, if the number of elements of L is finite, this follows from (L1)–(L3.)

‡ A sublattice A of L is any subsystem such that $a \in A$ and $a' \in A$ imply $a \cap a' \in A$ and $(a, a') \in A$.

§ That is, in which the subscripts run through transfinite ordinals.

‖ By the "sublattice generated by" is meant the least sublattice containing.

ber of elements in the various B_i^*. Finally, if the B_i are B-lattices, then so† is B^*.

Conversely, let B be any B-lattice, and let $B_1, \cdots, B_n$ be any finite or transfinite sequence of strongly coprime sublattices of B such that any $b \in B$ can be expressed as the meet $(b_{i_1}, \cdots, b_{i_m})$ of a finite number of $b_{i_k} \in B_{i_k}$.

For any $b \in B$ and $b' \in B$ we can evidently so reorder the B_i that $b = (b_1, \cdots, b_m)$, $b' = (b_1', \cdots, b_m')$, and $b \cap b' = b'' = (b_1'', \cdots, b_m'')$, where $b_i \in B_i$, $b_i' \in B_i$, $b_i'' \in B_i$, and m is finite. But by (L2)–(L3), we have

$$(b, b') = ((b_1, \cdots, b_m), (b_1', \cdots, b_m')) = ((b_1, b_1'), \cdots, (b_m, b_m')).$$

Further if we set $a_i = (b_i, b_i', b_i'')$, then

$$a_i \cap b'' = a_i \cap b \cap a_i \cap b' = a_i \cap (b_1, \cdots, b_m) \cap a_i \cap (b_1', \cdots, b_m'),$$

whence by (L5), setting $c_i = (b_1, \cdots, b_{i-1}, b_{i+1}, \cdots, b_m)$, and c_i' and c_i'' equal to the corresponding dashed expressions, we have

$$(a_i \cap c_i'', b_i'') = (a_i \cap c_i, b_i) \cap (a_i \cap c_i', b_i'),$$

whence, by strong coprimeness, after reduction, $b_i'' = b_i \cap b_i'$.

That is, B is a homeomorphic image of the subdirect product $B^* = B_1 \hat{x} \cdots \hat{x} B_n$. But if $b_i = b_i'$, and b^* in B^* is the image of (b_i, b_i') of B_i, then $b^* \cap [b_1, \cdots, b_n] = b_i \neq b_i' = b^* \cap [b_1', \cdots, b_n']$, whence, by (L1), $[b_1, \cdots, b_n] \neq [b_1', \cdots, b_n']$, and the homeomorphism is an isomorphism. In summary, we have proved the following theorem.

THEOREM 1. *A given B-lattice B (with largest element) is isomorphic with the subdirect product $B_1^* \cdots B_n^*$ (B_i^* any B-lattice with largest element) if and only if B contains strongly coprime sublattices $B_1, \cdots, B_n$ respectively isomorphic with $B_1^*, \cdots, B_n^*$ such that any $b \in B$ can be expressed as a meet $(b_{i_1}, \cdots, B_{i_m})$, where m is finite and $b_{i_k} \in B_{i_k}$.*

Notice that if n is finite, then a subdirect product is a direct product; while if $n = 2$, then strong coprimeness is equivalent to coprimeness.

4. *Uniqueness Theory.* Let L be any lattice (with a largest element), and suppose that L is isomorphic with two subdirect products $A_1 \hat{x} \cdots \hat{x} A_m$ and $B_1 \hat{x} \cdots \hat{x} B_n$. We know by the sec-

† The identical relations (L2)–(L5) can be checked seriatim.

ond paragraph of § 3, how to identify the A_i (and B_i) with strongly coprime sublattices of L in such a way that any element of L can be represented as the meet of a finite number of elements of the various A_i (or B_i). The reader can easily check the statement that, since $(a_i, a_j) = b_k$, $(b_k \in B_k)$, if and only if $a_i \in B_k$ and $a_j \in B_k$, each B_i is the subdirect product of its intersections with the various A_i; this proves the following statement.

THEOREM 2. *If* $L = A_1 \hat{x} \cdots \hat{x} A_m = B_1 \hat{x} \cdots \hat{x} B_n$ *is any lattice,*† *then* $L = F_{1,1} \hat{x} \cdots \hat{x} F_{m,n}$, *where* $A_i = F_{i,1} \hat{x} \cdots \hat{x} F_{i,n}$ *and* $B_j = F_{1,j} \hat{x} \cdots \hat{x} F_{m,j}$.

COROLLARY 1. *A lattice has at most one expression as a subdirect product of factors not themselves subdirect products.*

COROLLARY 2. *A finite lattice has a unique expression as the direct product of lattices not themselves direct products of lattices with fewer elements. The factors of any expression of the lattice as a direct product are direct products of the factors of this special decomposition into prime factors.*

These corollaries are of extremely general application.‡ We now assume in addition that L satisfies the following postulate.

(ϕ) Any sequence $a_1, a_2, a_3, \cdots$ of elements of L, such that $a_k < a_{k+1}$ for every k, is finite.

Well-order the expressions $L = L_1^i \hat{x} \cdots \hat{x} L_n^i$ of L as a subdirect product, and apply Theorem 2 iteratedly. If we concentrate our attention on the corresponding well-ordered set of meets $(a_1^i, \cdots, a_{\alpha_i}^i) = a$ representing a fixed $a \in L$ (each a_k^i lying in just one of the L_k^j for each $j \leqq i$, by Theorem 2), we see that the expression $(a_1^i, \cdots, a_{\alpha_i}^i)$ undergoes§ in virtue of (ϕ) at most a finite number of transmutations. Hence we can proceed through limit-numbers, and, by transfinite induction, we have the following result.

---

† By definition of subdirect product, either $m = n = 1$ and the theorem is trivial, or the A_i, B_j, and L have largest elements.

‡ See Theorem 3.1 of the author's paper *On the combination of subalgebras*, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464. This article will be cited in future references as "Subalgebras."

§ Each transmutation replaces an a_k^i by the meet of $a_{k'}^{i+1} > a_k^i$ and $a_{k''}^{i+1} > a_k^i$.

THEOREM 3. *A lattice satisfying* (ϕ) *has one and only one expression as a subdirect product of factors not themselves subdirect products.*

Theorem 3 can evidently be applied to the ideals in rings which satisfy the ideal-chain theorem.

5. *Standard Exceptions to* (L6). Let B be any B-lattice, suppose g_1, g_2, and g_3 to be any three elements of B, and refer to Tables I–III of "Subalgebras"—only replacing A_i, B_i, M_i, N_i, C_i, F_i, and H_i by a_i, b_i, m_i, n_i, c_i, f_i, and h_i.

Suppose $c_i = c_j$ for some $i \neq j$. Then $a = (c_i, c_j) = c_i \cap c_j = b$, whence $(g_1, h_1) = (g_1, h_1, h_2, h_3) = (g_1, f_1 \cap f_2 \cap f_3) = (f_2 \cap f_3) \cap (g_1, f_1)$ [by (L5)] $= f_2 \cap f_3 \cap f = f_2 \cap f_3$, which is to say, $(g_1, g_2, \cap g_3) = (g_1, g_2) \cap (g_1, g_3)$. If therefore (L6) is violated at all, we must have some instance where the c_k are all distinct, yet $(c_i, c_j) = a$ and $c_i \cap c_j = b$ for $i \neq j$, whence $(c_1, c_2 \cap c_3) \neq (c_1, c_2) \cap (c_1, c_3)$. This proves the following fact.

THEOREM 4. *If a B-lattice is not a C-lattice, it contains a sublattice of order five and fixed structure not a C-lattice.*

Combining Theorem 4 with the result, due to Dedekind,† that any lattice not a B-lattice contains a sublattice of order five and fixed structure not a B-lattice, we get the following result.

COROLLARY. *If a lattice is not a C-lattice, it contains a sublattice of order five which is not a C-lattice.*

6. *Specialization by Induction.* Suppose B of §5 satisfies condition (ϕ) of §4, and consider the exception referred to in Theorem 4. We can by (ϕ) choose $c_1^* \supset c_1$ *covered* by b (see §2). Theorems 8.1 and 9.1 of "*Subalgebras*" show us successively that c_3 covers (c_1^*, c_3), $b = c_2 \cap c_3$ covers $c_2^* = c_2 \cap (c_1^*, c_3)$, hence c_1^* and c_2^* both cover $a^* = (c_1^*, c_2^*)$. Similarly $b = c_3 \cap c_2^*$ covers $c_3^* = c_3 \cap (c_1^*, c_2^*)$, and, since $c_3^* \supset a^*$, $(c_1^*, c_3^*) = (c_2^*, c_3^*) = a^*$. This proves the following theorem.

THEOREM 5. *If B is any B-lattice satisfying* (ϕ), *then either B is a C-lattice or we can find a sublattice of B consisting of a least element* a^*, $c_1^* \neq c_2^* \neq c_3^* \neq c_1^*$ *covering* a^*, *and* $b = c_1^* \cap c_2^* = c_2^* \cap c_3^* = c_3^* \cap c_1^*$ *covering* c_1^*, c_2^*, *and* c_3^*.

† *Gesammelte Werke*, 1931, vol. II, p. 255.

7. *Facts about Ideals.* Throughout, R will be understood to denote a commutative ring which has a principal unit l and satisfies the Basis Theorem. Our notation will be that of van der Waerden† except that we shall denote by (A, B) the l.c.m., and by $A \cap B$ the g.c.f., of any two given ideals A and B. This is the inverse of van der Waerden's notation.

The following are either known or immediate corollaries of known results:

(1). *The only ideals in R are R and 0 if, and only if, R is a field.*

(2). *If I is a largest ideal in R, then $0:I$ is a least ideal if and only if it is a principal ideal.*

(3). *Any ideal I covered by R is a prime ideal.*

8. *Application of Theorem* 1. On the basis of Theorem 1, it is possible to reconstruct the combinatorial theory of an important class of ideals.

By an ideal *of genus* 1 will be meant any ideal I which contains an appropriate finite product $P_1^{n_1} \cdots P_\omega^{n_\omega}$ (where P_i denotes any ideal covered by R). We shall prove the following result.

THEOREM 6. *The ideals of genus* 1 *in R are a B-lattice, which is the subdirect product of the sublattices $\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3, \cdots$ of the primary ideals under the ideals $P_1, P_2, P_3, \cdots$ covered by R.*

That they are a lattice of which the $\mathfrak{P}_i$ are sublattices is obvious, while that they are a B-lattice follows from Theorem 27.1 of "*Subalgebras.*"

But the $\mathfrak{P}_i$ are strongly coprime, since if $Q_1, \cdots, Q_m$ satisfy the relation $Q_k \not\subseteq P_i$ for every k, and Q is primary under P_i, then $Q \cap (Q_1, \cdots, Q_m) = R$, being contained in no ideal covered by R. And by a theorem of E. Noether, any ideal can be expressed as the meet of a finite number of primary ideals. Theorem 6 is now merely a translation of Theorem 1 in terms of ideals.

9. *Application of Theorem* 5. It is not difficult to show from known results the following theorem.

THEOREM 7. *If R contains a largest ideal I, and another ideal $A \subset I$ for which $(A:I)/A$ is not a principal ideal, then the ideals of R are not a C-lattice.*

† *Moderne Algebra*, 1930–31; especially vol. 2, Chap 12, in which will be found the Basis and Ideal-chain Theorems.

For since l, commutativity, and the Basis Theorem are preserved under homeomorphism, we can assume $A=0$; while by (2) we can assume $(0:I)$ is not a least ideal.

By the Ideal-chain Theorem we can further choose a largest subideal $J>0$ in $0:I$, and then $x \notin J$, $y \notin Rx$ satisfying $y \in J$, and $w=x+y$. But a second homeomorphism permits us to assume $(Rx, Ry)=0$, yet $x \neq 0$, $y \neq 0$. This makes it obvious that $(Rx, Ry) \cap Rw \neq (Rx \cap Rw, Ry \cap Rw) \ni x$.

Conversely, suppose the ideals of R are not a C-lattice. By Theorem 5, R has a homeomorphic image R^* which contains three least ideals $A \neq B \neq C$ such that $A \cap B = B \cap C = C \cap A$.

Consider $R^*/(0:A)$; it is a field, whence, by (1), $0:A$, and similarly $0:B$ and $0:C$, are largest ideals† in R. For if $ra \neq 0$ $[a \in A, r \in R^*]$, then $ra \in A$ generates A; consequently r^{-1} exists such that $r^{-1}ra=a$ and $r^{-1}r \equiv l\ (0:A)$.

Again, if $0 \neq b \in B \subset A \cap C$, then $b=a+c$, where (since $a=0$ or $b=0$ would imply $B=C$ or $B=A$) $a \neq 0$, $c \neq 0$. And since $(A, C)=0$, $0=rb=r(a+c)=ra+rc$ implies $ra=rc=0$. Consequently $0:B \subset (0:A, 0:C)$, and $0:A=0:B=0:C=I$, where I is a largest ideal in R^*, yet, by (2), $0:I$ is not a principal ideal in R^*. Referring back to the corresponding ideals in R, we see that R does not satisfy the conclusions of Theorem 7.

We can combine Theorem 7, its converse, and Theorem 25.2 of "*Subalgebras*" in the following theorem.

THEOREM 8. *For the ideals of R to be isomorphic (with respect to l.c.m. and g.c.f.) with a system of point sets (with respect to sum and product), it is necessary and sufficient that if I is any largest ideal in R, and $A \subset I$ another ideal, then $(A:I)/A$ is a principal ideal in R/A.*

It is a corollary that the identity $A:(A:Q)=(Q \cap A)$ upon ideals is a sufficient condition for distributive combination.

HARVARD UNIVERSITY

† $R=0:A$ is of course excluded since $l \notin 0:A$.

A METRICAL PROPERTY OF POINT-SET TRANSFORMATIONS*

BY DEANE MONTGOMERY†

It is the purpose of this note to prove a theorem concerning distances between points of a metric space‡ under transformations of the space into itself. The distance between two points p and q of a metric space is denoted by $d(p, q)$. The point into which a point p is taken by n applications of a transformation T is denoted by p_n, that is, T takes p to p_1, p_1 to p_2, and so on. If $d(p, q)$ is less than, greater than, or equal to $d(p_1, q_1)$, T is said, respectively, to increase, decrease, or to leave invariant the distance between p and q. The theorem of this note shows that any continuous transformation of a suitably restricted metric space into itself leaves invariant the distance between some two distinct points of the space. In particular, it follows from the theorem that any continuous transformation of a closed or open n-sphere into itself leaves the distance between two distinct points of the sphere invariant. The theorem also applies to spherical surfaces.

It should be observed that in the following lemma the transformation is not assumed to be continuous as it is in the theorem. In both cases the phrase "into itself" means that the space goes into the whole of itself and not into a proper subset of itself.

LEMMA. *If E is a conditionally compact metric space and if T is any one-to-one transformation of E into itself which increases the distance between two points of E, then T must also decrease the distance between two points of E.*

Let p and q be two points of E whose distance is increased by T, that is, $d(p, q) < d(p_1, q_1)$. The proof will be made by contradiction. Suppose that the lemma is false and that, for every two points x and y of E, we have $d(x, y) \leqq d(x_1, y_1)$.

* Presented to the Society, March 30, 1934.

† National Research Fellow.

‡ For definitions of metric spaces and terms concerning them see Hausdorff, *Mengenlehre*, 1927.

Assume first that neither p nor q is a periodic point of the transformation. Since E is conditionally compact, the sequence $(p, p_1, p_2, \cdots)$ must have a subsequence $(p_{k_1}, p_{k_2}, p_{k_3}, \cdots)$ which satisfies the Cauchy convergence criterion. The sequence $(q_{k_1}, q_{k_2}, q_{k_3}, \cdots)$ must have a subsequence $(q_{n_1}, q_{n_2}, q_{n_3}, \cdots)$ which satisfies the Cauchy convergence criterion. Since $(n_1, n_2, n_3, \cdots)$ is a subsequence of $(p_1, p_2, p_3, \cdots)$, the sequence $(p_{n_1}, p_{n_2}, p_{n_3}, \cdots)$ also satisfies the convergence criterion. Therefore, for any positive number ϵ there is an N such that for $n_i > N$ and $n_j > N$,

$$d(p_{n_i}, p_{n_j}) < \epsilon, \quad \text{and} \quad d(q_{n_i}, q_{n_j}) < \epsilon.$$

Choose n_i and n_j so that $n_j > n_i$. Then according to the supposition that the lemma is false and to the above inequalities,

$$d(p, p_{n_j - n_i}) \leqq d(p_{n_i}, p_{n_j}) < \epsilon,$$

and

$$d(q, q_{n_j - n_i}) \leqq d(q_{n_i}, q_{n_j}) < \epsilon.$$

Let $n_j - n_i$ be denoted by r_1. It can be seen that there is a sequence $(r_1, r_2, r_3, \cdots)$ such that $(p_{r_1}, p_{r_2}, p_{r_3}, \cdots)$ converges to p and $(q_{r_1}, q_{r_2}, q_{r_3}, \cdots)$ converges to q. Hence for any ϵ, the integer i may be so chosen that $|d(p_{r_i}, q_{r_i}) - d(p, q)| < \epsilon$. In particular, there is some i such that $|d(p_{r_i}, q_{r_i}) - d(p, q)| < d(p_1, q_1) - d(p, q)$. This contradicts the fact that $d(p_1, q_1) \leqq d(p_{r_i}, q_{r_i})$. Therefore the supposition that the lemma was false was incorrect and there are two points x and y such that $d(x, y) > d(x_1, y_1)$.

Consider next the case in which both p and q are periodic, p of period m and q of period n. For the product mn, $p_{mn} = p$ and $q_{mn} = q$. Therefore $d(p_{mn}, q_{mn}) = d(p, q) < d(p_1, q_1)$. If the distance between no two points is decreased, it follows that $d(p_1, q_1) \leqq d(p_{mn}, q_{mn})$. Hence the distance between some two points must decrease.

The last case is that in which one of the points, say p, is periodic of period m and q is not periodic. The procedure in this case is much the same as in the first case. From the sequence $(q_m, q_{2m}, q_{3m}, \cdots)$ choose a subsequence $(q_{k_1}, q_{k_2}, q_{k_3}, \cdots)$ satisfying the Cauchy convergence criterion. As before it is possible to obtain a sequence $(q_{r_1}, q_{r_2}, q_{r_3}, \cdots)$ which converges

to q. For all n, r_n is a multiple of m, because $r_n = k_j - k_i$ and k_i and k_j are multiples of m. Therefore for all n, $p_{r_n} = p$, and $d(p_{r_i}, q_{r_i})$ converges to $d(p, q)$ and a contradiction may be obtained as before.

COROLLARY. *If E is a conditionally compact metric space and if T is a one-to-one transformation of E into itself which does not leave all distances invariant, then T must increase at least one distance and decrease at least one distance.*

If T increases some distance, the conclusion follows immediately from the lemma. If T decreases some distance, the lemma may be applied to the inverse of T and the conclusion obtained in that manner.

A metric space is said to be arc-wise connected if every two points in the space may be joined by a simple arc which lies in the space. A simple arc is defined to be any set which is homeomorphic to the set of real numbers on the closed interval from zero to one. In the following theorem it is assumed that the space contains at least two points.

THEOREM. *If E is an arc-wise connected conditionally compact metric space and if T is a one-to-one continuous transformation of E into itself, then there are two distinct points of E whose distance is invariant under T.*

If the distance between every pair of points is invariant, the theorem is obviously true. It may therefore be assumed that T alters the distance between some pair of points. It follows from the corollary that there are two points P and Q such that $d(P, Q) < d(P_1, Q_1)$ and two points R and S such that $d(R, S) > d(R_1, S_1)$. If P and R are coincident, Q and S can not be coincident, and if Q and S are coincident, P and R are not coincident. The notation may be so chosen that it is P and R which are not coincident. Then the points Q and S may or may not be coincident. There is an arc PR joining P and R and lying in E. It is necessary to consider several cases and subcases.

CASE I. The points Q and S are on PR and are coincident.

If Q and S are coincident, the two must be distinct from P and from R. A homeomorphism may be established between PQ and QR in which P corresponds to Q and Q corresponds to

R. If x is any point of PQ, its corresponding point on QR is denoted by x'. An essential fact is that x and x' are never coincident. Define a real-valued function $f(x)$ for each point x of PQ as follows:

$$f(x) = d(x, x') - d(x_1, x_1').$$

Since $f(P) = d(P, Q) - d(P_1, Q_1)$, $f(P)$ is negative. Since $f(Q) = d(Q, R) - d(Q_1, R_1) = d(S, R) - d(S_1, R_1)$, $f(Q)$ is positive. The function $f(x)$ is continuous and therefore for some point M of PR, $f(M) = d(M, M') - d(M_1, M_1') = 0$, and M and M' are the two distinct points desired.

CASE II. The points Q and S are on PR and are distinct.

Assume first that Q coincides with R. Then S can not coincide with P. Define a function $h(x)$ for all points x of arc PS in the following manner:

$$h(x) = d(x, R) - d(x_1, R_1).$$

The function $h(x)$ is continuous and, since $h(P)$ is negative and $h(S)$ is positive, there must be some point M on PS such that $h(M) = 0$. The points M and R are then the required points.

If S coincides with P, Q can not coincide with R and the procedure is the same as in the situation just completed.

If neither Q nor S coincides with either R or P, there are again two possibilities. As the first possibility assume that S separates P and Q on the arc PR. In this case a homeomorphism may be set up between the arc PS and the arc QR in which P corresponds to Q, S corresponds to R, and in which no two corresponding points coincide. The proof may be completed as in Case I by a function $f(x)$ defined on PS. As the second possibility Q may separate P and S on the arc PR. By proceeding with care it is also possible in this case to set up a homeomorphism between PS and QR in which P corresponds to Q, S corresponds to R, and in which no corresponding points coincide. Again the proof may be completed by means of a function $f(x)$ defined on PS.

CASE III. Either Q or S is not on PR.

As there is no difference between the two cases assume that

Q is not on PR. If Q and S are coincident, define a function $h(x)$ on PR as follows:

$$h(x) = d(x, Q) - d(x_1, Q_1).$$

This function is continuous; it is negative at P and positive at R and the proof is as before.

Suppose that Q and S are distinct and denote by QS the arc joining Q and S. If QS includes both P and R, the situation is essentially the same as that in which the arc PR includes all four points. It is therefore assumed that QS does not contain both P and R. Consider first the case in which QS does not include R. There is then a point A on PR which is not on QS and which is such that no point of PR between A and R is on QS. There is a point B on QS such that B is not on PR and such that no point of QS between Q and B is on PR. If a homeomorphism between PR and QS is set up in such a way that P corresponds to Q, A to B, and R to S, then no two corresponding points are coincident. Therefore a function $f(x)$ may be used as in previous cases. The points A and B are so chosen that $A \neq R$ and $B \neq Q$.

The remaining possibility is that QS does not include P. If S is not on PR, this reduces to the case just considered, where the parts played by the various points are interchanged. Assume, then, that S is on PR. Let K be the first point on QS which is a point of PR, the ordering on QS being assumed to begin at Q and continue to S. If K and R are coincident, use may be made of a homeomorphism between PS and QK. No two points of such a homeomorphism are coincident. If K is not R, form an arc C by adding the subarc QK of QS to the subarc KS of PR. In the event that K and S are coincident, KS is not a subarc but merely a point. This does not alter that which is to follow. The arc C joins Q and S, and since C does not contain R, the situation is reduced to one already treated.

HARVARD UNIVERSITY

CYCLIC FIELDS OF DEGREE p^n OVER F OF CHARACTERISTIC p*

BY A. A. ALBERT

1. *Introduction.* The theory of cyclic fields is a most interesting chapter in the study of the algebraic extensions of an abstract field F. When F is a modular field of characteristic p, a prime, particular attention is focussed on the case of cyclic fields Z of degree p^n over F. Such fields of degree p, p^2 were determined by E. Artin and O. Schreier.†

In the present paper I shall give *a determination of all cyclic fields Z of degree p^n over F of characteristic p.*

2. *Normed Equations.* An equation

$$\lambda^p = \lambda + a, \qquad (a \text{ in } F), \tag{1}$$

is called a *normed equation.* If x is any root of (1), then so are $x+1, x+2, \cdots, x+p-1$. Using this fact, Artin-Schreier have proved the following lemmas.

LEMMA 1. *A normed equation is either cyclic or has all of its roots in F. Every cyclic field of degree p over F may be generated by a root of a normed equation.*

LEMMA 2. *Let $F(x)$ be cyclic of degree p over F,*

$$x^p = x + a, \qquad (a \text{ in } F). \tag{2}$$

Then a quantity y of $F(x)$ which is not in F satisfies a normed equation if and only if

$$y = kx + b, \quad (k = 1, 2, \cdots, p;\ b \text{ in } F). \tag{3}$$

LEMMA 3. *Let c in Z have degree $t \leqq p-2$ in x. Then there exists a quantity $g = g(x)$ in Z such that*

$$g(x + 1) - g(x) = c. \tag{4}$$

Moreover, g is uniquely determined up to an additive constant in F.

---

* Presented to the Society, March 30, 1934.

† Hamburg Abhandlungen, vol. 5 (1926–7), pp. 225–231.

By applying Lemmas 1, 2, 3, Artin-Schreier proved the following fact.

LEMMA 4. *Every cyclic field Z of degree p over F is the sub-field of cyclic overfields Z_2 of degree p^2 over F. If $Z_1 = F(x_1)$, $x_1^p = x_1 + a_1$, a_1 in F, then all such fields Z_2 are obtained by*

$$Z_2 = F(x_2), \qquad x_2^p = x_2 + a_2, \qquad (a_2 \text{ in } Z_1), \tag{5}$$

where a_2 ranges over all solutions of (4) in the case

$$c = (x_1 + a_1)^{p-1} - x_1^{p-1}. \tag{6}$$

A generating automorphism S of Z_2 is given by

$$x_1^S = x_1 + 1, \qquad x_2^S = x_2 + x_1^{p-1}, \tag{7}$$

so that

$$x_2^{S^\nu} = x_2 + x_1^{p-1} + (x_1 + 1)^{p-1} + \cdots + (x_1 + \nu - 1)^{p-1}, \qquad (\nu = 1, 2, \cdots), \tag{8}$$

and in particular

$$x_2^{S^p} = x_2 + x_1^{p-1} + \cdots + (x_1 + p - 1)^{p-1} = x_2 - 1. \tag{9}$$

As an immediate corollary of (9) we have the following lemma.

LEMMA 5. *Let $Z = F(x)$, $x^p = x + a$, be cyclic of degree p over F. Then*

$$T_{Z|F}(x^{p-1}) \equiv x^{p-1} + (x + 1)^{p-1} + \cdots + (x + p - 1)^{p-1} = -1. \tag{10}$$

3. *Generating Automorphisms.* Now let $Z = Z_n$ be any cyclic field of degree p^n over F and let S be a generating automorphism of the cyclic automorphism group of Z. It is well known that

$$Z_n > Z_{n-1} > \cdots > Z_1 > Z_0 = F, \tag{11}$$

where Z_i is uniquely determined, is cyclic of degree p^i over F, cyclic of degree p over Z_{i-1}. Moreover the automorphism S applied in Z_i may be taken as generating the automorphism group of Z_i with

$$Q_i = S^{p^{n-i}} \tag{12}$$

as identity automorphism for Z_i. In fact Z_i is defined as the set of all quantities of Z_n (and no others) unaltered by the automorphism Q_i.

We may consider Z_i as cyclic of degree p over Z_{i-1}. Then the group of Z_i over Z_{i-1} is evidently generated by Q_{i-1},

$$(Q_{i-1})^p = Q_i. \tag{13}$$

If b_i is any quantity of Z_i, then we write

$$T_{Z_i|F}(b_i) = b_i + b_i S + \cdots + b_i S^{p^i-1}. \tag{14}$$

We then evidently have

$$T_{Z_i|F}(b_i) = T_{Z_{i-1}|F}\left[T_{Z_i|Z_{i-1}}(b_i)\right], \tag{15}$$

where

$$T_{Z_i|Z_{i-1}}(b_i) = b_i + b_i Q_{i-1} + \cdots + b_i Q_{i-1}^{p-1} \tag{16}$$

is evidently in Z_{i-1}.

The field Z_i is cyclic of degree p over Z_{i-1} so that, by Lemma 1,

$$Z_i = Z_{i-1}(x_i), \qquad x_i^p = x_i + a_i, \qquad (a_i \text{ in } Z_{i-1}). \tag{17}$$

Moreover, x_i is not in Z_{i-1}, so that $F(x_i)$ is in Z_i but not in Z_{i-1}. The cyclic field Z_{i-1} contains every proper sub-field of Z_i and hence must contain $F(x_i)$, if $F(x_i)$ is a proper sub-field of Z_i. Thus, in fact, we have

$$Z_i = F(x_i). \tag{18}$$

We may now prove the following fact.

LEMMA 6. *Let* $b_{i+1} = (x_1 x_2 \cdots x_i)^{p-1} = x_i^{p-1} b_i$. *Then* b_{i+1} *is in* Z_i *and*

$$T_{Z_i|F}(b_{i+1}) = (-1)^i. \tag{19}$$

For b_i is in Z_{i-1} and is unaltered by the automorphism Q_{i-1}. Hence $T_{Z_i|Z_{i-1}}(b_{i+1}) = b_i T_{Z_i|Z_{i-1}}(x_i^{p-1})$. Since Q_{i-1} is a generating automorphism of Z_i over Z_{i-1}, some power S_i of Q_{i-1} carries x_i into $x_i + 1$. But then Lemma 5 implies $T_{Z_i|Z_{i-1}}(x_i^{p-1}) = -1$. Hence

$$T_{Z_i|F}(b_{i+1}) = T_{Z_{i-1}|F}\left[b_i T_{Z_i|Z_{i-1}}(x_i^{p-1})\right] = -T_{Z_{i-1}|F}(b_i). \tag{20}$$

By repeated application of this recursion formula, we evidently obtain (19).

Let S be a generating automorphism of Z_n. Then

$$(x_i S)^p = x_i S + a_i S.$$

But evidently x_i^S is a primitive quantity of Z_i of degree p over Z_{i-1}, so that, by Lemma 2,

$$x_i S = k_i x_i + b_i, \qquad (k_i = 1, 2, \cdots, p-1;\ b_i \text{ in } Z_{i-1}). \tag{21}$$

Then $x_i^{S^2} = k_i x_i^S + b_i^S = k_i^2 x_i + b_{i2}$, and finally $x_i^{S^p} = k_i^p x_i + b_{ip}$. Hence, if $m = p^n$, we have $x_i^{S^m} = k_i^m x_i + b_{im} = x_i$. But then $k_i^m = 1$. Since $k_i^p = k_i$, we evidently have $k_i^m = k_i = 1$. Thus

$$x_i S = x_i + b_i, \qquad (i = 1, 2, \cdots, n). \tag{22}$$

Moreover,

$$(x_i^S)^p = x_i^p + b_i^p = x_i + a_i + b_i^p = x_i + b_i + a_i S,$$

$$a_i S - a_i = b_i^p - b_i. \tag{23}$$

The automorphism Q_{i-1} is a generating automorphism of Z_i over Z_{i-1} and replaces x_i by $x_i + h_i$, $(h_i = 1, 2, \cdots, p-1)$. But

$$x_i^{Q_{i-1}} = x_i + T_{Z_{i-1}|F}(b_i) = x_i + h_i,$$

so that

$$T_{Z_{i-1}|F}(b_i) = h_i \neq 0 \qquad (h_i = 1, 2, \cdots, p-1;\ i = 1, \cdots, n). \tag{24}$$

Conversely, let b_i satisfy (24), a_i be determined by (23), and let $x_1^p = x_1 + a_1$ be irreducible in F. Then $Z_n = F(x_n)$ is cyclic of degree p^n over F when Z_n is defined by (17), $Z_i = F(x_i)$, and S generates the automorphism group of Z_n. For assume this true for $Z_1, Z_2, \cdots, Z_{n-1}$, and define $Z_n = Z_{n-1}(x_n)$. The degree of Z_n over F is then p^n, for otherwise x_n is in Z_{n-1}, by Lemma 1, and hence $(x_n)^{Q_{n-1}} = x_n$, contrary to (24) and (22). Moreover, (22) defines an automorphism S of Z_n which has order at least p^{n-1}, since S generates the automorphism group of Z_{n-1}. But S actually has order p^n, since $Q_{n-1} = S^{p^{n-1}}$ alters x_n. Hence the group of automorphisms of Z_n has a cyclic sub-group of order p^n, the degree of Z_n, and Z_n is cyclic. It follows that $Z_n = F(x_n)$. We have proved the following result.

LEMMA 7. *Every cyclic field of degree p^n over F is generated by a quantity x_n such that*

$$x_i^p = x_i + a_i, \qquad a_i \text{ in } Z_{i-1} = F(x_{i-1}), \quad (i = 1, 2, \cdots, n), \tag{25}$$

and $x_1^p = x_1 + a_1$ is irreducible in F. If S is a generating automorphism of the group of Z_n, then

$$x_i^S = x_i + b_i, \qquad T_{Z_i|F}(b_{i+1}) = h_i, \quad (h_i = 1, \cdots, p-1;\ i = 1, \cdots, n), \tag{26}$$

with

$$b_i^p - b_i = a_i^S - a_i, \qquad (i = 2, \cdots, n). \tag{27}$$

Conversely, every field Z_n defined by (25), (26), (27) *and $x_1^p = x_1 + a$ irreducible in F, is cyclic of degree p^n over F with generating automorphism S given by* (26).

Let c_i be an arbitrary quantity of Z_i and write

$$c_i = \sum_{j_r = 0, 1, \cdots, p-1} \lambda_{j_1 j_2 \cdots j_i} x_1^{j_1} x_2^{j_2} \cdots x_i^{j_i}$$

with coefficients λ in F. If $\lambda_{p-1,\ p-1, \cdots, p-1} = 0$, we call c_i a non-maximal quantity of Z_i. We may prove the following lemma.

LEMMA 8. *If $b_i = (x_1 x_2 \cdots x_{i-1})^{p-1}$, the polynomials*

$$c_{i-1} = b_i^p - b_i = [(x_1 + a_1)(x_2 + a_2) \cdots (x_{i-1} + a_{i-1})]^{p-1} - (x_1 x_2 \cdots x_{i-1})^{p-1}, \qquad (i = 2, \cdots, n), \tag{28}$$

are non-maximal and (27) *have solutions a_i in Z_{i-1} which are unique up to an arbitrary additive constant in F. Then the a_i define cyclic fields Z_i, $(i = 2, \cdots, n)$, containing Z_1, where Z_i is cyclic of degree p^i over F. In fact, if c_i is any non-maximal quantity of Z_i, there exist solutions d_i in Z_i of*

$$c_i = d_i^S - d_i, \qquad d_i^S \equiv d_i(x_1 + b_1, \cdots, x_i + b_i), \quad (i = 1, \cdots, n), \tag{29}$$

which are unique up to an additive constant in F.

For evidently $T_{Z_i|F}(b_{i+1}) = (-1)^i = h_i \neq 0$, so that $(26)_2$ are satisfied. It is thus sufficient to prove the existence of the a_i satisfying (27) and hence sufficient to prove the existence and uniqueness of solutions of (29) of which (27) is a special case.

We know that Lemma 8 is true for $n=1, 2$* by Lemmas 3, 4. Hence assume Lemma 8 true in its entirety for Z of degree $p, p^2, \cdots, p^{i-1}$. Then, by our assumption (29), there *exists* a Z_i of degree p^i over F, the equation (29) has a unique solution in Z_{i-1}, and we wish to prove (29) also has a unique solution in Z_i and hence the existence of Z_{i+1}.

Write

$$c_i = \lambda_t x_i^t + \cdots + \lambda_0, \qquad (\lambda_j \text{ in } Z_{i-1}).$$

If λ_t is a non-maximal quantity of Z_{i-1}, then, by our above assumption, $\lambda_t = \mu_t^S - \mu_t$, ($\mu_t$ in Z_{i-1}). But then

$$(\mu_t x_i^t)^S - \mu_t x_i^t = \mu_t^S (x_i + b_i)^t - \mu_t x_i^t = \lambda_t x_i^t + \cdots,$$

so that $c_i - [(\mu_t x_i^t)^S - (\mu_t x_i^t)]$ has degree at most $t-1$.

If λ_t is maximal, then $t < p-1$ and c_i has leading term

$$\lambda(x_1 x_2 \cdots x_{i-1})^{p-1} x_i^t = \lambda b_i x_i^t, \qquad \lambda \neq 0 \text{ in } F.$$

But then $t+1 \neq 0$,

$$\lambda(t+1)^{-1}[(x_i^{t+1})^S - x_i^{t+1}] = \lambda(t+1)^{-1}[(x_i + b_i)^{t+1} - x_i^{t+1}]$$
$$= \lambda b_i x_i^t + \cdots,$$

so that $c_i - \{[\lambda(t+1)^{-1} x_i^{t+1}]^S - [\lambda(t+1)^{-1} x_i^{t+1}]\}$ has degree at most t in x_i and non-maximal leading coefficient. A repeated application of the above process may evidently be made to obtain a quantity δ_i in Z_i such that $c_i - (\delta_i^S - \delta_i) = \gamma_{i0}$, ($\gamma_{i0}$ in Z_{i-1}). But γ_{i0} may be taken non-maximal as above with $t=0, t+1=1$, and hence

$$\gamma_{i0} = \gamma_i^S - \gamma_i, \qquad c_i = d_i^S - d_i, \qquad d_i = \delta_i + \gamma_i.$$

Now let $c_i = d_i^S - d_i = d_{i0}^S - d_{i0}$. Then $(d_{i0} - d_i)^S = d_{i0}^S - d_i^S = d_{i0} - d_i$. The only quantities of Z_i unaltered by S are quantities of F. Hence $d_{i0} - d_i = \lambda$ in F. We have proved Lemma 8. We shall now prove our principal theorem.

THEOREM. *Every cyclic field Z_1 of degree p over F of characteristic p is the sub-field of cyclic overfields of degree p^n. Write*

$$Z_1 = F(x_1), \qquad x_1^p = x_1 + a_1, \qquad (a_1 \text{ in } F). \tag{30}$$

* Note that (28) is true for $n=2$ by Lemma 3, is vacuous for $n=1$. Hence Z_2 is defined by Lemma 4. This is the *first* step in our induction, the case $i=2$.

Then all such fields Z_n are given by

$$(31)\quad Z_i = F(x_i),\quad x_i^p = x_i + a_i,\quad a_i \text{ in } F(x_{i-1}),\quad (i = 2, \cdots, n),$$

where a_i is the unique (up to an arbitrary additive constant in F) solution of

$$(32)\quad \begin{aligned} &a_i(x_1 + b_1, \cdots, x_{i-1} + b_{i-1}) - a_i \\ &\quad = [(x_1 + a_1) \cdots (x_{i-1} + a_{i-1})]^{p-1} - (x_1 x_2 \cdots x_{i-1})^{p-1}, \\ &b_i \equiv (x_1 x_2 \cdots x_{i-1})^{p-1}, \qquad (i = 2, \cdots, n). \end{aligned}$$

Conversely, all fields defined by (30), (31), (32) *with $x_1^p = x_1 + a_1$ irreducible in F are cyclic of degree p^n with generating automorphism S given by*

$$(33)\quad x_0 = 1,\quad x_i^S = x_i + (x_0 x_1 x_2 \cdots x_{i-1})^{p-1},\quad (i = 1, \cdots, n).$$

For we have proved that the fields defined above are cyclic, in Lemma 8. Assume now, conversely, that Z_n is cyclic of degree p^n over F and that we have proved the above result for its subfields $Z_1, \cdots, Z_{n-1}$. Let $x_n^S = x_n + d_n$ by Lemma 7 and write $d_n = \beta b_n + g_n$, where $b_n = (x_1 x_2 \cdots x_{i-1})^{p-1}$, β is in F, and $-g_n$ is a non-maximal polynomial in Z_{n-1}. By Lemma 8, we have also

$$-g_n = h_n^S - h_n, \qquad (h_n \text{ in } Z_{n-1}).$$

We then let $y_n = x_n + h_n$, so that

$$y_n^S = x_n^S + h_n^S = x_n + \beta b_n + g_n + h_n - g_n = y_n + \beta b_n.$$

Moreover, $Z_n = F(x_n) = F(y_n)$, since it is evident that y_n generates $Z_{n-1}(x_n)$ over Z_{n-1} and hence also $F(x_n)$, by Lemma 7 (in which we proved $Z_{n-1}(x_n) = F(x_n)$).

But now we have shown that we may take $d_n = \beta b_n$ without loss of generality. Since

$$T_{Z_{n-1}|F}(d_n) = \beta T_{Z_{n-1}|F}(b_n) = (-1)^{n-1}\beta = k_n, \qquad (k_n = 1, \cdots, p-1),$$

the quantity β is a non-zero integer. There exists an integer γ such that $\gamma\beta = 1$ and, if we write $z_n = \gamma x_n$, we have $z_n^S = \gamma x_n^S = \gamma(x_n + \beta b_n) = z_n + b_n$. Evidently $F(x_n) = F(z_n)$ while z_n satisfies (33). By Lemma 7, (27), we have also (32) for $i = n$; and we have proved the theorem.

THE INSTITUTE FOR ADVANCED STUDY

ON THE EXPANSION COEFFICIENTS OF THE FUNCTIONS $u/\mathrm{sn}u$ AND u^2/sn^2u †

BY R. D. JAMES‡

1. *Introduction.* It is well known that the elliptic function $\mathrm{sn}(u, k)$ may be expanded in the form

$$\mathrm{sn}(u, k) = \sum_{n=0}^{\infty} \frac{u^n}{n!} S_n(k^2),$$

where

$$(1)\quad S_0(k^2) = 0, \quad S_1(k^2) = 1, \quad S_n(k^2) = \sum_{r=0}^{[(n-1)/2]} s_r(n) k^{2r}, \quad (n \geqq 2).$$

Moreover, $s_r(2j) = 0$ for all values of r and j, and $s_r(2j+1)$ is an integer >0 for $0 \leqq r \leqq j$. Similarly we have

$$(2)\quad \frac{u}{\mathrm{sn}(u, k)} = \sum_{n=0}^{\infty} \frac{u^n}{n!} G_n(k^2),$$

where

$$G_0(k^2) = 1, \qquad G_n(k^2) = \sum_{r=0}^{[n/2]} g_r(n) k^{2r}, \qquad (n \geqq 1),$$

with $g_r(2j+1) = 0$ for all values of r and j, and $g_r(2j) \neq 0$ for $0 \leqq r \leqq j$. In particular

$$(3)\quad g_0(2) = g_1(2) = 1/3.$$

Again,

$$(4)\quad \frac{u^2}{\mathrm{sn}^2(u, k)} = \frac{1}{2} k^2 u^2 + \sum_{n=0}^{\infty} \frac{u^n}{n!} T_n(k^2),$$

where

$$T_0(k^2) = 1, \qquad T_n(k^2) = \sum_{r=0}^{[n/2]} t_r(n) k^{2r}, \qquad (n \geqq 1),$$

with $t_r(2j+1) = 0$ for all r and j, and $t_r(2j) \neq 0$ for $0 \leqq r \leqq j$. In particular,

$$(5)\quad t_0(2) = 2/3, \quad t_1(2) = -1/3, \quad t_0(4) = -t_1(4) = t_2(4) = 8/5.$$

---

† Presented to the Society, March 30, 1934.

‡ National Research Fellow.

The values of $s_r(n)$, $g_r(n)$ and $t_r(n)$ for $r=0, 1, 2, 3$, were given without proof by Hermite.† A proof for $s_r(n)$, for $r=0, 1, 2, 3$, was given by Gruder‡ but no proof for $g_r(n)$ and $t_r(n)$ seems to have been published. In this paper we shall prove the following results:

$$g_0(n) = -i^n(2^n-2)B_n; \tag{6}$$

$$g_1(n) = i^n 2^{-n}n(2(2^n-2)B_n - 1 - (-1)^n); \tag{7}$$

$$t_0(n) = -i^n(n-1)2^nB_n; \tag{8}$$

$$t_1(n) = i^n n(n-1)2^{n-2}B_n; \tag{9}$$

$$\begin{aligned} t_2(2) &= 0, \\ t_2(n) &= -i^n n(n-1)((2n-7)2^{n-6}B_n - 2^{n-8}(1+(-1)^n)), \end{aligned} \tag{10}$$

where $n \neq 2$. The B_n are the Bernoulli numbers defined by the equation

$$\frac{1}{2}u \operatorname{ctn} \frac{1}{2}u = \sum_{n=0}^{\infty} \frac{u^n}{n!} i^n B_n. \tag{11}$$

The proof depends on a number of recursion formulas for the Bernoulli numbers and the method may be applied to calculate further values of $g_r(n)$ and $t_r(n)$.

2. *Recursion Formulas for the Bernoulli Numbers.* To simplify the writing of formulas we adopt the notation

$$(f+g)^n = \sum\sum_{\substack{\alpha+\beta=n \\ \alpha,\beta \geqq 0}} \frac{n!}{\alpha!\beta!} f_\alpha g_\beta,$$

$$(f(*)+g(*))^n = \sum\sum_{\substack{\alpha+\beta=n \\ \alpha,\beta \geqq 0}} \frac{n!}{\alpha!\beta!} f(\alpha)g(\beta),$$

and similarly for $(f+g+h)^n$ and $(f(*)+g(*)+h(*))^n$. Let

$$\begin{aligned} A_n &= (n-1)B_n, & J_n &= 2^nB_n, & K_n &= (n-1)2^nB_n, \\ M_n &= n(n-1)2^nB_n, & P_n &= n(n-1)(2n-7)2^nB_n, \\ R_n &= (2^n-2)B_n, & U_n &= (n-1)(2^n-2)B_n, \\ V_n &= \tfrac{1}{2}n(1+(-1)^n), & W_n &= n(n-1)2^{n-1}(1+(-1)^n). \end{aligned} \tag{12}$$

† Collected Works, vol. 3, pp. 236–237.

‡ Wiener Sitzungsberichte, vol. 126, IIa (1917).

The results which we shall require are as follows:

$$(13)\qquad (B+B)^n = \begin{cases} -(n-1)B_n, & (n \neq 2), \\ -(n-1)B_n + \tfrac{1}{2}, & (n=2); \end{cases}$$

$$(14)\qquad 4(B+B+B)^n = 2(n-1)(n-2)B_n + n(n-1)B_{n-2};$$

$$(15)\qquad (R+R)^n = -(n-1)2^n B_n = -K_n;$$

$$(16)\qquad 2(R+R+R)^n = (n-1)(n-2)R_n - n(n-1)R_{n-2};$$

$$(17)\qquad (J+J)^n = \begin{cases} -(n-1)J_n, & (n \neq 2), \\ -(n-1)J_n + 2, & (n=2); \end{cases}$$

$$(18)\qquad (B+J)^n = \begin{cases} -(n-1)B_n, & (n \neq 2), \\ -(n-1)B_n + 1, & (n=2); \end{cases}$$

$$(19)\qquad 6(A+A)^n = -(n-2)(n-3)A_n + n(n-1)A_{n-2};$$

$$(20)\qquad 6(K+K)^n = -(n-2)(n-3)K_n + 4n(n-1)K_{n-2};$$

$$(21)\qquad \begin{aligned} 6(A+K)^n &= -(n-2)(n-3)A_n \\ &\quad + n(n-1)(3\cdot 2^{n-3}+1)A_{n-2}; \end{aligned}$$

$$(22)\qquad \begin{aligned} -6(R+R+U)^n &= 6(K+U)^n \\ &= -(n-2)(n-3)U_n + n(n-1)U_{n-2}; \end{aligned}$$

$$(23)\qquad -(R+R+V)^n = (K+V)^n = n(n-1)R_{n-2};$$

$$(24)\qquad (K+W) = \begin{cases} 4n(n-1)K_{n-2}, & (n \neq 4), \\ 4n(n-1)K_{n-2} - 192, & (n=4); \end{cases}$$

$$(25)\qquad \begin{aligned} 6(K+P)^n + 24(M+M)^n + 6(P+K)^n \\ = -(n-2)(n-3)P_n + 4n(n-1)P_{n-2} \\ + 32n(n-1)M_{n-2} - 24n(n-1)K_{n-2}. \end{aligned}$$

These formulas are readily proved and in most cases we merely indicate the method of proof.

PROOF OF (13). The function $x=(u/2)\operatorname{ctn}(u/2)$ satisfies the differential equations

$$(26)\qquad 4u\frac{dx}{du} = -4x^2 + 4x - u^2,$$

$$(27)\qquad 2u^2\frac{d^2x}{du^2} - 4u\frac{dx}{du} = 4x^3 - 4x + u^2x.$$

It follows from (11) and (26) that

$$4\sum_{n=1}^{\infty}\frac{u^n}{(n-1)!}i^nB_n = -4\sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(B+B)^n + 4\sum_{n=0}^{\infty}\frac{u^n}{n!}i^nB_n + 2\frac{u^2}{2!}i^2.$$

Hence

$$8B_2 = -4(B+B)^2 + 4B_2 + 2,$$
$$4nB_n = -4(B+B)^n + 4B_n, \qquad (n \neq 2),$$

which proves (13). The formula (14) follows in the same way from (11) and (27).

PROOF OF (15). It is known that

$$-u\csc u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^nR_n.$$

Since $u^2\csc^2 u = u^2 + u^2\operatorname{ctn}^2 u$, we have

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(R+R)^n = -2\frac{u^2}{2!}i^2 + \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n2^n(B+B)^n,$$

$$(R+R)^n = \begin{cases} 2^n(B+B)^n = -(n-1)2^nB_n, & (n \neq 2), \\ -2 + 2^2(B+B)^2 = -(2-1)2^2B_2, & (n = 2). \end{cases}$$

PROOF OF (16). The function $y = u\csc u$ satisfies the equation

$$u^2\frac{d^2y}{du^2} - 2u\frac{dy}{du} = 2y^3 - 2y - u^2y.$$

Hence

$$-n(n-1)R_n + 2nR_n = -2(R+R+R)^n + 2R_n - n(n-1)R_{n-2},$$
$$2(R+R+R)^n = (n-1)(n-2)R_n - n(n-1)R_{n-2}.$$

PROOF OF (17). In (13) replace B_n by $2^{-n}J_n$.

PROOF OF (18). We have

$$\frac{1}{2}u^2\operatorname{ctn}\frac{1}{2}u\operatorname{ctn}u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(B+J)^n,$$

and

$$\frac{1}{2}u^2 \operatorname{ctn}\frac{1}{2}u \operatorname{ctn} u = \frac{1}{4}u^2 \operatorname{ctn}\frac{1}{2}u - \frac{1}{4}u^2$$

$$= \frac{1}{2}\frac{u^2}{2!}i^2 + \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(B+B)^n.$$

Hence

$$(B+J)^n = \begin{cases} (B+B)^n = -(n-1)B_n, & (n \neq 2), \\ \frac{1}{2} + (B+B)^2 = -(n-1)B_n + 1, & (n = 2). \end{cases}$$

PROOF OF (19). The function

$$z = \frac{1}{4}u^2 \csc^2\frac{1}{2}u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n 2^{-n}(R+R)^n = -\sum_{n=0}^{\infty}\frac{u^n}{n!}i^n A_n$$

satisfies the differential equation

$$u^2\frac{d^2z}{du^2} - 4u\frac{dz}{du} = 6z^2 - 6z - u^2 z.$$

Hence

$$-n(n-1)A_n + 4nA_n = 6(A+A)^n + 6n - n(n-1)A_{n-2}.$$

PROOF OF (20). In (19) replace A_n by $2^{-n}K_n$.

PROOF OF (21). We have

$$\frac{1}{4}u^4 \csc^2\frac{1}{2}u \csc^2 u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(A+K)^n,$$

and

$$\frac{1}{4}u^4 \csc^2\frac{1}{2}u \csc^2 u = \frac{1}{16}u^4 \csc^4\frac{1}{2}u + \frac{1}{4}u^4 \csc^2 u$$

$$= \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(A+A)^n - \frac{1}{4}\sum_{n=0}^{\infty}\frac{u^{n+2}}{n!}i^n 2^n A_n.$$

Hence

$$6(A+K)^n = 6(A+A)^n + 3\cdot 2^{n-3}n(n-1)A_{n-2}.$$

PROOF OF (22). The fact that $-(R+R+U)^n = (K+U)^n$ follows from (15). Moreover, $U_n = K_n - 2A_n$. Hence, from (20), (21), and (12),

$$\begin{aligned} 6(K+U)^n &= 6(K+K)^n - 12(A+K)^n \\ &= -(n-2)(n-3)K_n + 4n(n-1)K_{n-2} \\ &\quad + 2(n-2)(n-3)A_n \\ &\quad - 2n(n-1)(3\cdot 2^{n-3}+1)A_{n-2} \\ &= -(n-2)(n-3)U_n + n(n-1)U_{n-2}. \end{aligned}$$

PROOF OF (23). As before, $-(R+R+V)^n=(K+V)^n$. Moreover, we have

$$u \sin u = -\sum_{n=0}^{\infty} \frac{u^n}{n!}\frac{1}{2}n(i^n+(-i)^n) = -\sum_{n=0}^{\infty}\frac{u^n}{n!}i^nV_n.$$

Hence

$$u^3 \csc^2 u \sin u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(K+V)^n$$

and

$$u^3 \csc^2 u \sin u = u^3 \csc u = -\sum_{n=0}^{\infty}\frac{u^{n+2}}{n!}i^nR_n.$$

Then we have $(K+V)^n=n(n-1)R_{n-2}$.

PROOF OF (24). We have

$$\begin{aligned} 4u^2 \cos 2u &= -\sum_{n=0}^{\infty}\frac{u^n}{n!}n(n-1)2^{n-1}(i^n+(-i)^n) \\ &= -\sum_{n=0}^{\infty}\frac{u^n}{n!}i^nW_n. \end{aligned}$$

Then

$$4u^4 \csc^2 u \cos 2u = \sum_{n=0}^{\infty}\frac{u^n}{n!}i^n(K+W)^n,$$

and

$$\begin{aligned} 4u^4 \csc^2 u \cos 2u &= 4u^4 \csc^2 u - 8u^4 \\ &= -192\frac{u^4}{4!}i^4 - 4\sum_{n=0}^{\infty}\frac{u^{n+2}}{n!}i^nK_n. \end{aligned}$$

Hence

$$(K+W)^n = \begin{cases} 4n(n-1)K_{n-2}, & (n \neq 4), \\ 4n(n-1)K_{n-2} - 192, & (n = 4). \end{cases}$$

PROOF OF (25). Since

$$K_\alpha P_\beta + 4M_\alpha M_\beta + P_\alpha K_\beta = (\alpha+\beta)(2\alpha+2\beta-7)K_\alpha K_\beta,$$

we have, from (20) and (12),

$$\begin{aligned} &6(K+P)^n + 24(M+M)^n + 6(P+K)^n \\ &\quad = 6n(2n-7)(K+K)^n \\ &\quad = n(2n-7)\{-(n-2)(n-3)K_n + 4n(n-1)K_{n-2}\} \\ &\quad = -(n-2)(n-3)P_n + 4n(n-1)P_{n-2} \\ &\qquad + 32n(n-1)M_{n-2} - 24n(n-1)K_{n-2}. \end{aligned}$$

3. *The Proof of* (6) *and* (7). Since $\mathrm{sn}(u, 0) = \sin u$, we have, from (2) and (12),

$$\sum_{n=0}^{\infty} \frac{u^n}{n!} g_0(n) = -\sum_{n=0}^{\infty} \frac{u^n}{n!} i^n R_n,$$

$$g_0(n) = -i^n R_n = -i^n(2^n - 2)B_n.$$

Next, the function $w = \mathrm{sn}(u, k)$ satisfies† the equations

$$\left(\frac{dw}{du}\right)^2 = (1-w^2)(1-k^2w^2), \quad \frac{d^2w}{du^2} = -(1+k^2)w + 2k^2w^3.$$

It follows that $\xi = u/\mathrm{sn}(u, k) = u/w$ satisfies the equation

$$u^2\frac{d^2\xi}{du^2} - 2u\frac{d\xi}{du} = 2\xi^3 - 2\xi - (1+k^2)u^2\xi. \tag{28}$$

Hence, from (2),

$$\begin{aligned} n(n-1)G_n - 2nG_n &= 2(G+G+G)^n - 2G_n \\ &\quad - (1+k^2)n(n-1)G_{n-2}, \\ 2(G+G+G)^n &= (n-1)(n-2)G_n + (1+k^2)n(n-1)G_{n-2}, \\ 2\sum_{\substack{\lambda+\mu+\nu=r \\ \lambda,\mu,\nu\geq 0}}\sum\sum &(g_\lambda(*) + g_\mu(*) + g_\nu(*))^n \\ &= (n-1)(n-2)g_r(n) + n(n-1)(g_r(n-2) + g_{r-1}(n-2)). \end{aligned}$$

† Gruder, loc. cit.

When $r=1$, the last equation becomes

$$6(g_0(*)+g_0(*)+g_1(*))^n=(n-1)(n-2)g_1(n)+n(n-1)(g_1(n-2)+g_0(n-2)). \tag{29}$$

Now let

$$Z_n=i^n2^{-3}n(2(2^n-2)B_n-1-(-1)^n)=i^n2^{-2}(U_n+R_n-V_n)$$

and consider $6(g_0(*)+g_0(*)+Z)^n$. By (16), (22), and (23), this equals

$$\begin{aligned}
&i^n2^{-2}\{6(R+R+U)^n+6(R+R+R)^n-6(R+R+V)^n\}\\
&\quad=i^n2^{-2}\{(n-2)(n-3)U_n-n(n-1)U_{n-2}\\
&\qquad+3(n-1)(n-2)R_n-3n(n-1)R_{n-2}+6n(n-1)R_{n-2}\}\\
&\quad=i^n2^{-2}\{(n-1)(n-2)(U_n+R_n-V_n)\\
&\qquad-n(n-1)(U_{n-2}+R_{n-2}-V_{n-2})+4n(n-1)R_{n-2}\}\\
&\quad=(n-1)(n-2)Z_n+n(n-1)Z_{n-2}+n(n-1)g_0(n-2).
\end{aligned}$$

This proves that Z_n satisfies the recursion formula (29). Since, from (3), we have $g_1(2)=1/3=Z_2$, it follows that $g_1(n)=Z_n$ for all values of $n\geqq 2$. It is readily verified that $g_1(0)=Z_0=g_1(1)=Z_1=0$, and this completes the proof.

4. *The Proof of* (8), (9), (10). We have

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}t_0(n)=\frac{u^2}{\mathrm{sn}^2(u,0)}=u^2\csc^2 u=-\sum_{n=0}^{\infty}\frac{u^n}{n!}i^nK_n,$$

$$t_0(n)=-i^nK_n=-i^n(n-1)2^nB_n.$$

Next, we have†

$$(-1)^r t_r(2j)=\sum_{\nu=0}^{r}\binom{j-\nu}{j-r}t_\nu(2j),$$

$$t_1(2j)=-\tfrac{1}{2}jt_0(2j)=i^{2j}2^{-2}2j(2j-1)2^{2j}B_{2j}.$$

Since $t_1(2j+1)=B_{2j+1}=0$, this proves (9). Again, the function $\eta=u^2/\mathrm{sn}^2(u,k)-k^2u^2/2$ satisfies the equation

$$u^2\frac{d^2\eta}{du^2}-4u\frac{d\eta}{du}=6\eta^2-6\eta-2(2-k^2)u^2\eta-k^2\frac{1}{2}u^4.$$

† Gruder, loc. cit.

Hence, from (4),

$$n(n-1)T_n - 4nT_n = 6(T+T)^n - 6T_n - 2(2-k^2)n(n-1)T_{n-2}, \quad (n \neq 4),$$

$$6\sum_{\substack{\lambda+\mu=r \\ \lambda,\mu\geqq 0}} \sum (t_\lambda(*) + t_\mu(*))^n = (n-2)(n-3)t_r(n) + 4n(n-1)t_r(n-2) - 2n(n-1)t_{r-1}(n-2), \quad (n \neq 4).$$

When $r=2$, the last equation becomes

$$\begin{aligned} &6(t_0(*)+t_2(*))^n + 6(t_1(*)+t_1(*))^n + 6(t_2(*)+t_0(*))^n \\ &= (n-2)(n-3)t_2(n) + 4n(n-1)t_2(n-2) \\ &\quad - 2n(n-1)t_1(n-2), \qquad (n \neq 4). \end{aligned} \tag{30}$$

Let

$$\begin{aligned} H_n &= -i^n n(n-1)((2n-7)2^{n-6}B_n - 2^{n-8}(1+(-1)^n)) \\ &= -i^n 2^{-6}P_n + i^n 2^{-7}W_n, \qquad (n \neq 2). \\ H_2 &= -i^2 2^{-6}P_2 - i^2 2^{-7}W_2 = 0. \end{aligned}$$

and consider

$$6(t_0(*)+H)^n + 6(t_1(*)+t_1(*))^n + 6(H+t_0(*))^n \qquad (n \neq 4).$$

By (25), (24), and (12), this equals

$$\begin{aligned} &i^n 2^{-6}\{6(K+P)^n + 24(M+M)^n + 6(P+K)^n\} \\ &\quad - i^n 2^{-7}12\{(K+W)^n - n(n-1)K_{n-2}W_2\} \\ &= i^n 2^{-6}\{-(n-2)(n-3)P_n + 4n(n-1)P_{n-2} \\ &\quad + 32n(n-1)M_{n-2} - 24n(n-1)K_{n-2}\} \\ &\quad - i^n 2^{-7}12\{4n(n-1)K_{n-2} - 8n(n-1)K_{n-2}\} \\ &= i^n 2^{-6}\{-(n-2)(n-3)(P_n - 2^{-1}W_n) \\ &\quad + 4n(n-1)(P_{n-2} - 2^{-1}W_{n-2}) + 32n(n-1)M_{n-2}\} \\ &= (n-2)(n-3)H_n + 4n(n-1)H_{n-2} - 2n(n-1)t_1(n-2). \end{aligned}$$

This shows that H_n satisfies the recursion formula (30). Since $t_2(4) = 8/5 = H_4$, it follows that $t_2(n) = H_n$ for all $n \geqq 4$. Finally, we verify that $t_2(n) = H_n = 0$ for $n = 0, 1, 2, 3$. This completes the proof of (8), (9), and (10).

CAMBRIDGE UNIVERSITY

OFFICIAL COMMUNICATIONS

Meetings of the Society have been fixed at the following times and places:

WILLIAMSTOWN, MASSACHUSETTS, SUMMER MEETING AND COLLOQUIUM, September 4-7, 1934.

Abstracts must be in the hands of Associate Secretary J. R. Kline, 501 West 116th St., New York City, not later than August 1. Abstracts received by July 6 will appear in the July issue of the Bulletin. Professor Norbert Wiener will deliver a series of colloquium lectures which were planned by himself and the late Dr. R. E. A. C. Paley on *Topics from harmonic analysis in the complex domain.* By invitation of the Program Committee, Professor J. A. Shohat will deliver an address entitled *On the expansion of functions in series of orthogonal polynomials.*

NEW YORK CITY, October 27, 1934.

Abstracts must be in the hands of Associate Secretary J. R. Kline, 501 West 116th St., New York City, not later than September 29. Abstracts received by the Secretaries by September 7 will appear in the September issue of this Bulletin. By invitation of the Program Committee, Professor Emmy Noether, visiting lecturer at Bryn Mawr College, will deliver an address entitled *Modern hypercomplex series.*

LINCOLN, NEBRASKA, November 30-December 1, 1934.

Abstracts must be in the hands of Associate Secretary M. H. Ingraham, University of Wisconsin, Madison, Wisconsin, not later than November 3. All these will appear in the November issue of this Bulletin. By invitation of the Program Committee, addresses will be delivered by Professors G. A. Bliss and A. J. Kempner.

UNIVERSITY OF CALIFORNIA AT LOS ANGELES, December 1, 1934.

Abstracts must be in the hands of Associate Secretary T. M. Putnam, University of California, Berkeley, California, not later than November 3. All these will appear in the November issue of this Bulletin.

PITTSBURGH, PENNSYLVANIA, Annual Meeting, December 27-31, 1934.

Abstracts must be in the hands of Associate Secretary J. R. Kline, 501 West 116th St., New York City, not later than November 22, 1934. Abstracts received by the Secretaries by November 9 will appear in the November issue of this Bulletin.

R. G. D. RICHARDSON, *Secretary of the Society.*

Articles for insertion in the BULLETIN should be addressed to E. R. HEDRICK, Editor of the BULLETIN, University of California at Los Angeles. Reviews should be sent to W. R. LONGLEY, Yale University, New Haven, Conn. Notes should be sent to H. W. KUHN, Ohio State University, Columbus, Ohio.

Subscriptions to the BULLETIN, orders for back numbers, and inquiries in regard to non-delivery of current numbers should be addressed to the American Mathematical Society, 450-459 Ahnaip St., Menasha, Wis., or 501 West 116th St., New York.

The initiation fees and the annual dues of members of the Society (see this BULLETIN, p. 322, May, 1930; and the List of Officers and Members, September, 1932, p. 66), are payable to the Treasurer of the Society, Professor G. W. Mullins, 501 West 116th St., New York City.

WHOLE NUMBER 419

CONTENTS

PAGE

The Asymptotic Solutions of Ordinary Linear Differential Equations of the Second Order, with Special Reference to the Stokes Phenomenon. By R. E. LANGER 545

Composition of Arithmetical Identities. By E. T. BELL.. 583

Continued Fractions and Cross-Ratio Groups of Cremona Transformations. By H. S. WALL..... 587

On the Minimizing Property of the Harmonic Function. By E. J. MCSHANE..... 593

Simplification of the Whitehead-Huntington Set of Postulates for the Algebra of Logic. By A. H. DIAMOND.... 599

On Analytic Functions with Positive Imaginary Parts. By J. L. DOOB and B. O. KOOPMAN..... 601

Metrically Transitive Point Transformations. By M. H. MARTIN..... 606

On the Lattice Theory of Ideals. By GARRETT BIRKHOFF 613

A Metrical Property of Point-Set Transformations. By DEANE MONTGOMERY..... 620

Cyclic Fields of Degree p^n over F of Characteristic p*. By A. A. ALBERT..... 625

On the Expansion Coefficients of the Functions $u/\mathrm{sn}u$ and u^2/sn^2u. By R. D. JAMES..... 632

For official Communications and Notices, see the inside of the back cover.

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN